



US010000798B2

(12) **United States Patent**  
**Williams**

(10) **Patent No.:** **US 10,000,798 B2**  
(45) **Date of Patent:** **\*Jun. 19, 2018**

(54) **POLYMERASE-NUCLEIC ACID COMPLEX**

7,462,452 B2 12/2008 Williams et al.  
7,745,116 B2\* 6/2010 Williams ..... C12Q 1/6869  
435/6.12

(71) Applicant: **Pacific Biosciences of California, Inc.**,  
Menlo Park, CA (US)

8,592,148 B2 11/2013 Williams et al.  
2002/0039738 A1 4/2002 Williams et al.

(72) Inventor: **John G. K. Williams**, Lincolnville, ME  
(US)

2002/0042071 A1 4/2002 Williams et al.  
2002/0137047 A1 9/2002 Baskin et al.

(73) Assignee: **Pacific Biosciences of California, Inc.**,  
Menlo Park, CA (US)

2002/0168678 A1 11/2002 Williams et al.  
2003/0044781 A1 3/2003 Korlach et al.

(\*) Notice: Subject to any disclaimer, the term of this  
patent is extended or adjusted under 35  
U.S.C. 154(b) by 406 days.

2003/0134807 A1 7/2003 Hardin et al.  
2003/0143581 A1 7/2003 Franzen et al.

This patent is subject to a terminal dis-  
claimer.

2003/0143598 A1 7/2003 Garimella et al.  
2003/0165859 A1 9/2003 Nazarenko et al.

2003/0215816 A1 11/2003 Sundararajan et al.  
2004/0023248 A1 2/2004 O'Malley

2004/0067503 A1 4/2004 Tan et al.  
2004/0086914 A1 5/2004 Cole et al.  
2005/0042633 A1 2/2005 Williams

(21) Appl. No.: **14/509,006**

FOREIGN PATENT DOCUMENTS

(22) Filed: **Oct. 7, 2014**

WO 1999/049831 A2 10/1999  
WO 2004/020654 A2 3/2004

(65) **Prior Publication Data**

US 2015/0086994 A1 Mar. 26, 2015

OTHER PUBLICATIONS

**Related U.S. Application Data**

(63) Continuation of application No. 11/377,136, filed on  
Mar. 15, 2006, now Pat. No. 9,045,798, which is a  
continuation of application No. 11/118,031, filed on  
Apr. 29, 2005, now Pat. No. 7,462,452.

Bruck and O'Donnell. *Genome Biology* 2001; 2: 3001.1-3001.3.\*  
Tran et al. *The Journal of Biological Chemistry* 2012; 287: 39732-  
39741.\*

(60) Provisional application No. 60/567,202, filed on Apr.  
30, 2004.

Reddy et al. *Proceedings of the National Academy of Sciences*,  
USA 1993; 90: 3211-3215.\*

(51) **Int. Cl.**  
**C12Q 1/68** (2018.01)  
**C07H 21/04** (2006.01)

Tabor et al. *The Journal of Biological Chemistry* 1987; 262:  
16212-16223.\*

(52) **U.S. Cl.**  
CPC ..... **C12Q 1/6869** (2013.01); **C07H 21/04**  
(2013.01); **B01J 2219/005** (2013.01); **B01J**  
**2219/00497** (2013.01); **B01J 2219/00576**  
(2013.01); **B01J 2219/00648** (2013.01); **B01J**  
**2219/00722** (2013.01); **C12Q 1/6818**  
(2013.01); **C12Q 1/6825** (2013.01); **C12Q**  
**2525/113** (2013.01); **C12Q 2525/117**  
(2013.01); **C12Q 2563/107** (2013.01); **C12Q**  
**2563/149** (2013.01); **C12Q 2565/518**  
(2013.01)

Bedford et al. *Proceedings of the National Academy of Sciences*,  
USA 1997; 94: 479-484.\*

(58) **Field of Classification Search**  
None  
See application file for complete search history.

Johnson et al. *The Journal of Biological Chemistry* (papers in press)  
2003; M301366200v1. 54 pages.\*

(56) **References Cited**

U.S. PATENT DOCUMENTS

6,127,127 A 10/2000 Eckhardt et al.  
6,232,075 B1 5/2001 Williams  
6,255,083 B1 7/2001 Williams  
6,306,607 B2 10/2001 Williams  
6,316,229 B1 11/2001 Lizardi et al.  
6,361,944 B1 3/2002 Mirkin et al.  
6,548,264 B1 4/2003 Tan et al.  
6,610,499 B1 8/2003 Fulwyler et al.  
6,762,048 B2 7/2004 Williams  
6,869,764 B2 3/2005 Williams et al.

Eisenbrandt et al. *Nucleic Acids Research* 2002; 30: 1379-1386.\*  
Peliska et al. *Biochemistry* 1994; 33: 13817-13823.\*

Druillennec et al. *The Journal of Biological Chemistry* 1999; 274:  
11283-11288.\*

Williams et al. *Nucleic Acids Research* 2008; 36: e121.\*  
Huber et al. *The Journal of Biological Chemistry* 1987; 262:  
16224-16232.\*

Martinez-Rucobo et al. *The EMBO Journal* 2011; 30: 1302-1310.  
(Year: 2011).\*

(Continued)

*Primary Examiner* — Angela M. Bertagna  
(74) *Attorney, Agent, or Firm* — Kilpatrick Townsend &  
Stockton LLP

(57) **ABSTRACT**

The present invention provides novel compositions, meth-  
ods and apparatus for DNA sequencing that can be per-  
formed, e.g., in a two-electrode chamber. The present inven-  
tion also provides a method for sequencing a nucleic acid  
comprising immobilizing a plurality of complexes compris-  
ing a target nucleic acid, a primer nucleic acid, and a  
polymerase onto a surface, contacting the surface with a  
plurality of charged particles comprising a nucleotide phos-  
phate by applying an electric field, reversing the electric  
field to transport unbound charged particles away from the  
surface, and detecting the incorporation of a nucleotide  
phosphate into a single molecule of the primer nucleic acid.

**26 Claims, 23 Drawing Sheets**

(56)

## References Cited

## OTHER PUBLICATIONS

Sevostyanova et al. *Molecular Cell* 2011; 43: 253-262. (Year: 2011).\*

Mueser et al. *Virology Journal* 2010; 7: 359. (Year: 2010).\*

Wang et al. *Cell* 1997; 89: 1087-1099. (Year: 1997).\*

Anazawa, T. et al., "Electrophoretic Quantitation of Nucleic Acids without Amplification by Single-Molecule Imaging," *Analytical Chemistry*, Oct. 1, 2002, vol. 74, No. 19, pp. 5033-5038.

Braslavsky, I. et al., "Sequence information can be obtained from single DNA molecules," *PNAS*, Apr. 1, 2003, vol. 100, No. 7, pp. 3960-3964.

Curulli et al., "Bionzyme amperometric probes for choline and choline esters assembled with nonconducting electrosynthesized polymers," *Electroanalysis*, 2001, vol. 13, No. 3, pp. 236-242.

Femino, A.M. et al., "Visualization of Single RNA Transcripts in Situ," *Science*, Apr. 24, 1998, vol. 280, pp. 585-590.

Gerion et al., "Sorting fluorescent nanocrystals with DNA," *J Am Chem Soc* (2002) 124:7070-7074.

Heller et al., "Active microelectronic chip devices which utilize controlled electrophoretic fields for multiplex DNA hybridization and other genomic applications," *Electrophoresis*, 2000, vol. 21, pp. 157-164.

Kerman et al., "Electrochemical Coding of Single-Nucleotide Polymorphisms by Monobase-Modified Gold Nanoparticles," *Analytical Chemistry*, 2004, vol. 76, pp. 1877-1884.

Martinez-Zaguilan et al., "pH and Drug Resistance. I. Functional Expression of Plasmalemmal V-type H1-ATPase in Drug-Resistant

Human Breast Carcinoma Cell Lines," *Biochemical Pharmacology*, 1999, 57:1037-1046.

Niemeyer, C.M., "Self-assembled nanostructures based on DNA: towards the development of nanobiotechnology," *Current Opinion in Chemical Biology*, 2000, vol. 4, pp. 609-618.

Ronaghi, et al. "A sequencing method based on real-time pyrophosphate," *Science* (1998) 281(5375): 363 & 365.

Rybak et al., "Primary cell cultures from murine kidney and heart differ in endosomal pH," *Journal of Cell Physiology*, 1998, 176:216-222.

Shin et al., "A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron," *Nature*, 2004, vol. 427, No. 6975, pp. 618-621.

Sosnowski, et al. "Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control," *Proceedings of the National Academy of Science, USA* (1997) 94: 1119-1123.

Taylor, J.R. et al., "Probing Specific Sequences on Single DNA Molecules with Bioconjugated Fluorescent Nanoparticles," *Analytical Chemistry*, May 1, 2000, vol. 72, No. 9, pp. 1979-1986.

U.S. Appl. No. 11/377,136, filed Mar. 15, 2006 (allowed; unpublished); Inventor: Williams et al.; Confirmation No. 5054.

Yang et al., "Nanometer fluorescent hybrid silica particle as ultrasensitive and photostable biological labels," *Analyst*, 2003, vol. 128, pp. 462-466.

Zhao et al., "Ultrasensitive DNA Detection Using Highly Fluorescent Bioconjugated Nanoparticles", *Journal of the American Chemical Society*, vol. 125, No. 38, Aug. 2003, pp. 11474-11475.

\* cited by examiner

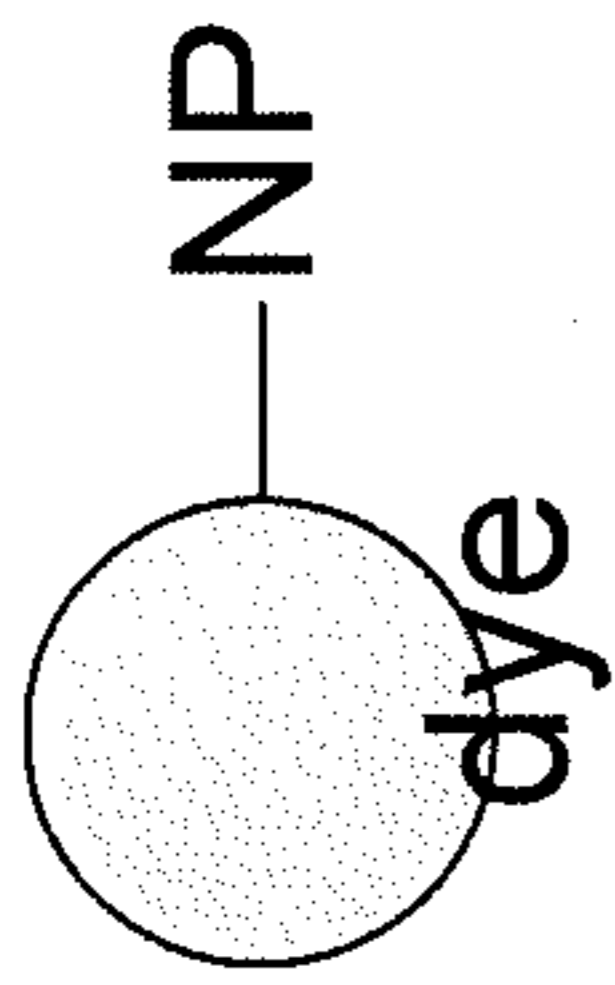


FIG. 1A

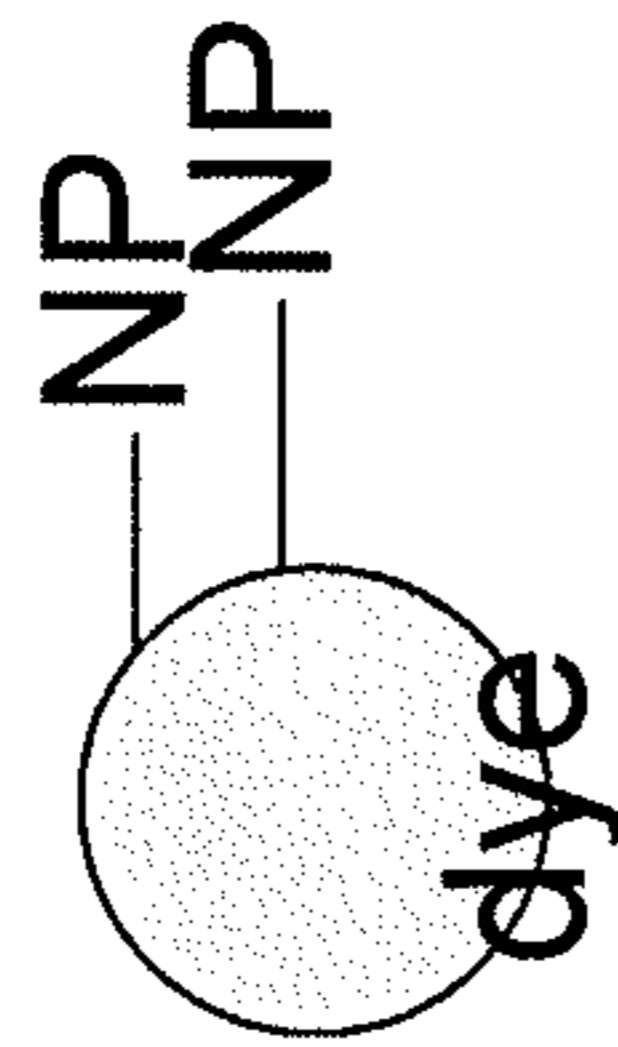


FIG. 1B

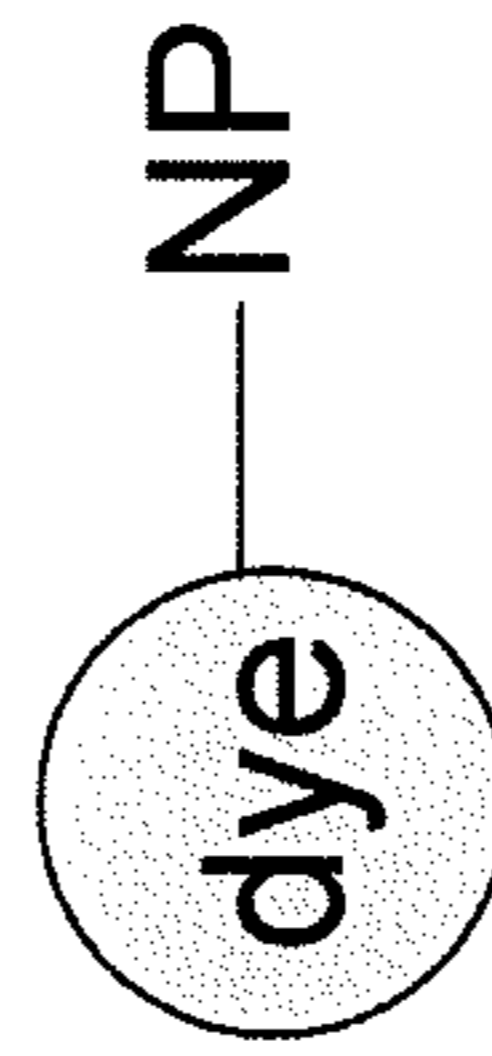


FIG. 1C

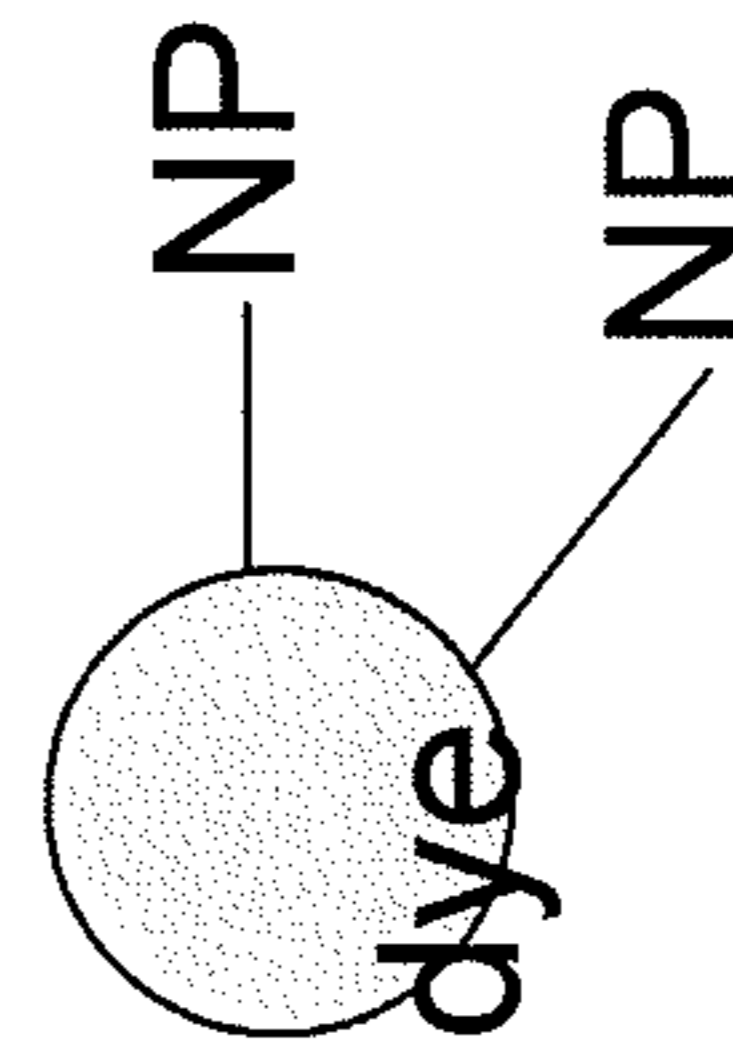


FIG. 1D

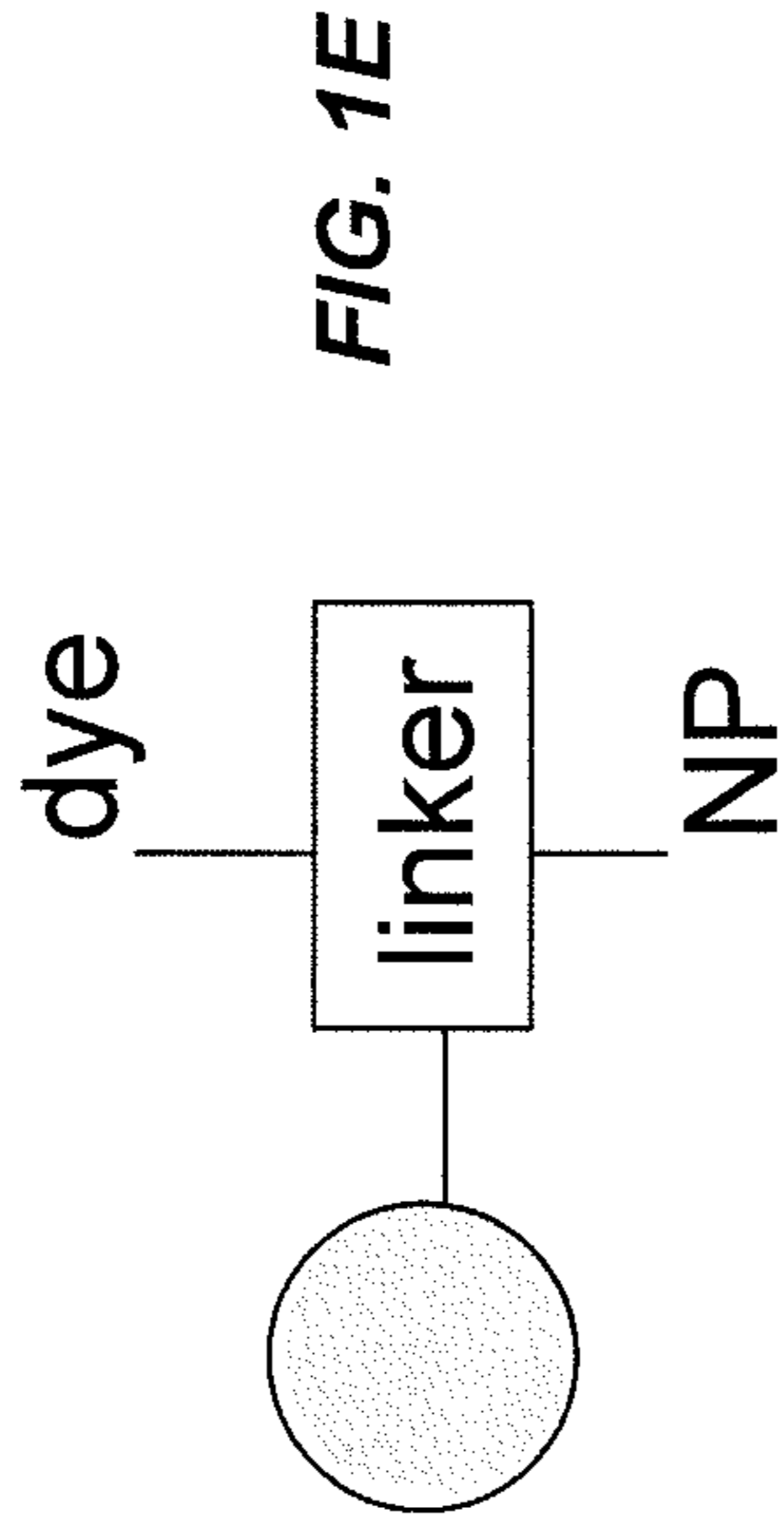


FIG. 1E

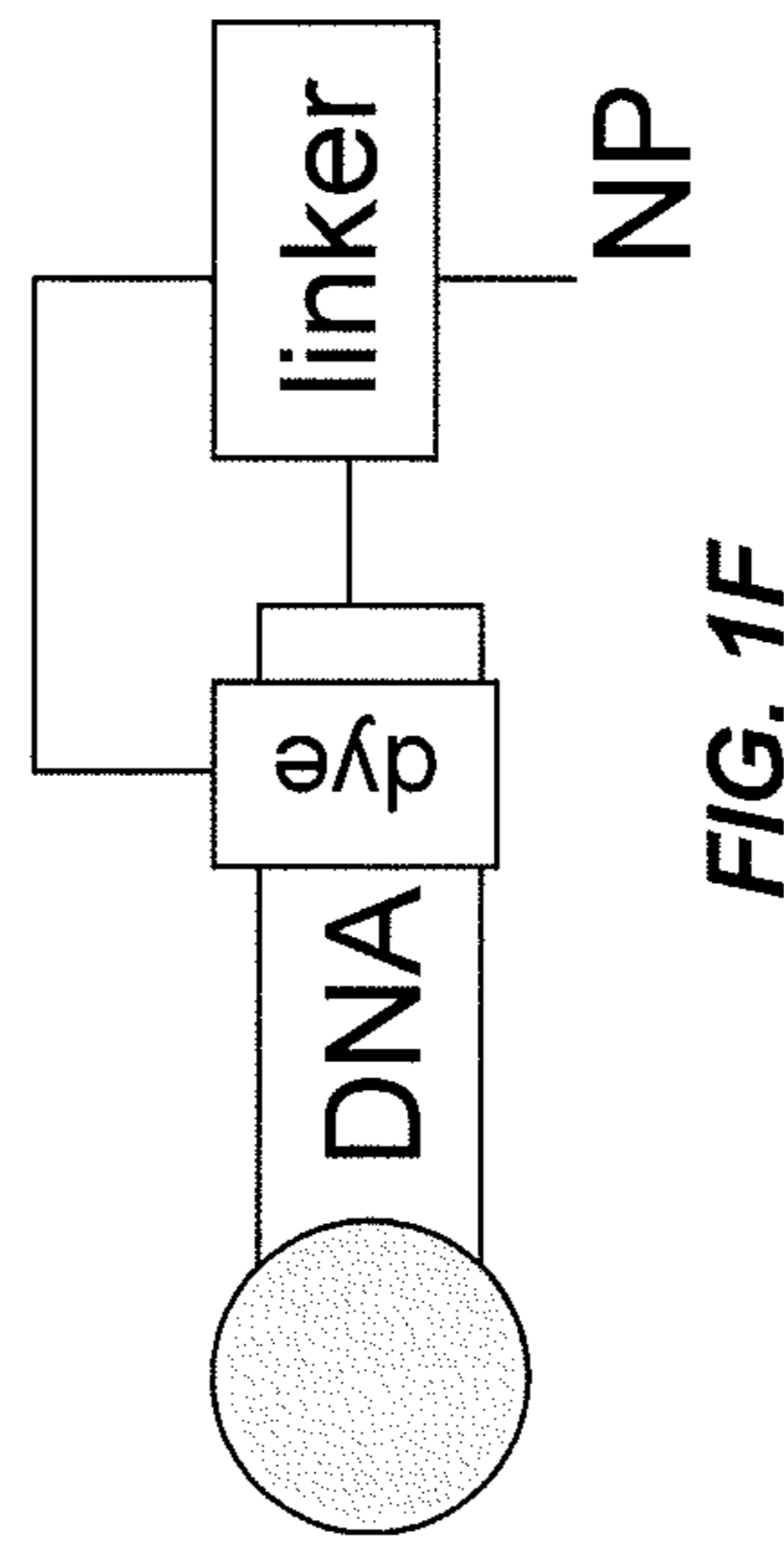


FIG. 1F

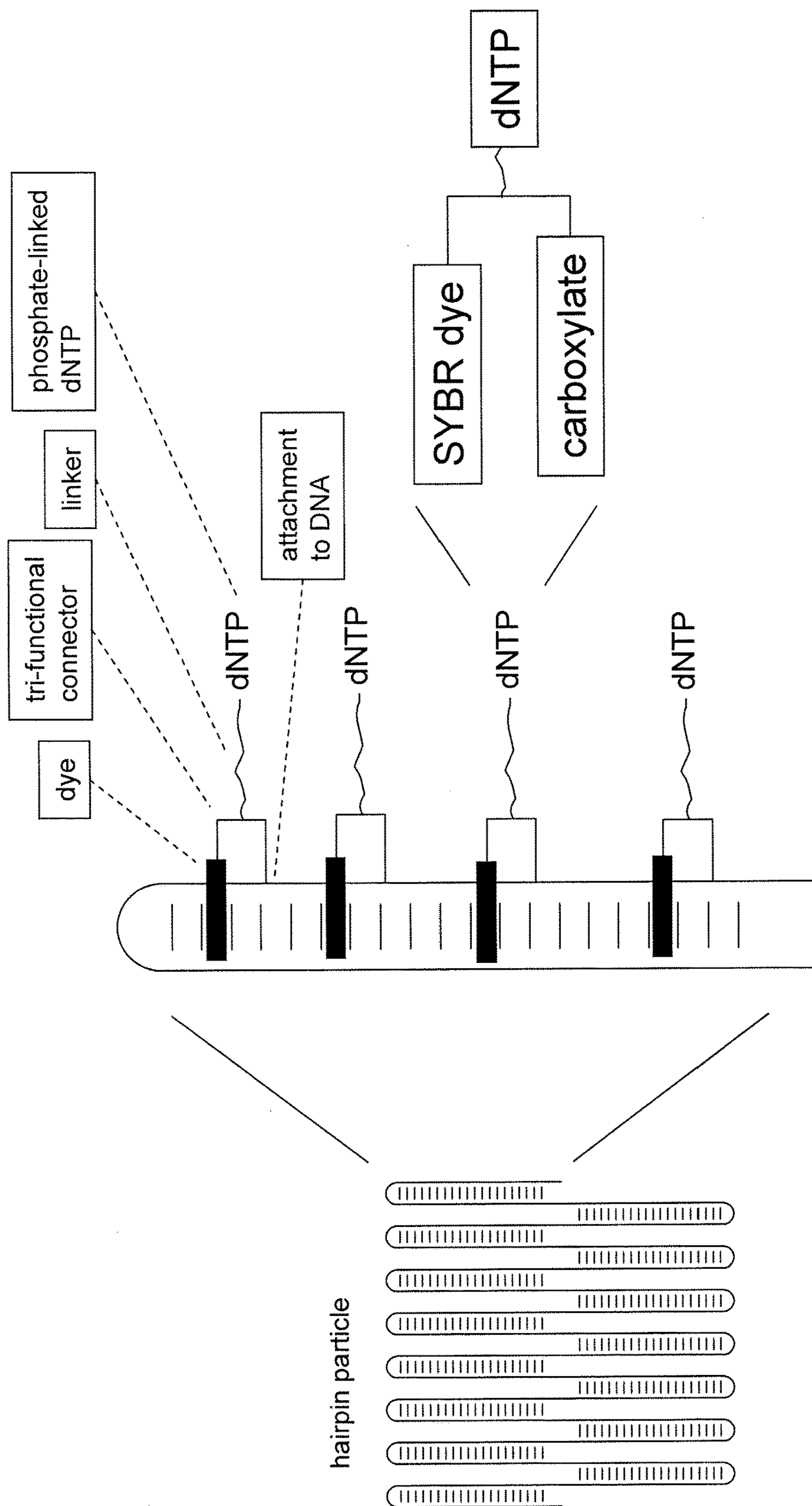
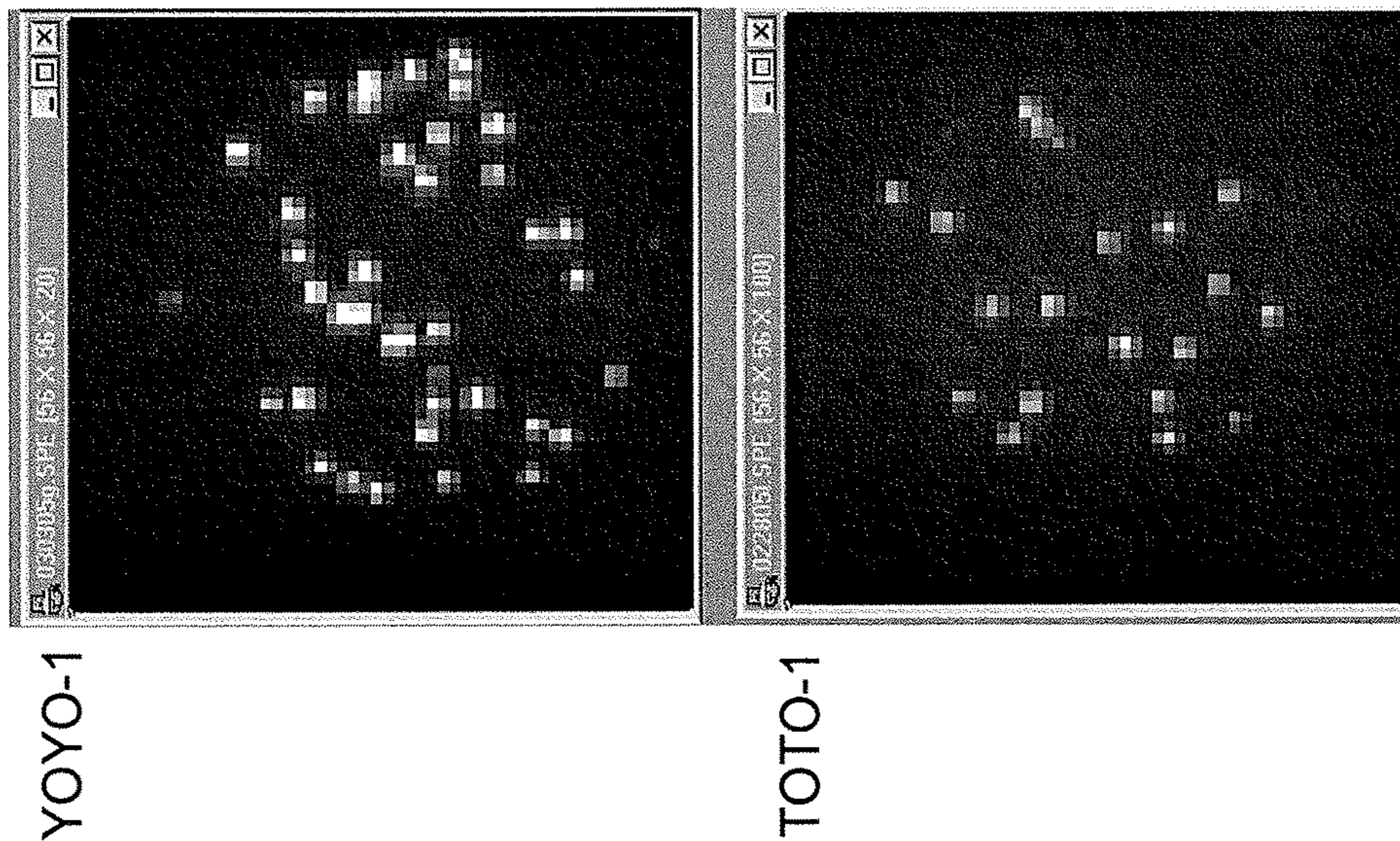


FIG. 2B

FIG. 2C

FIG. 2A



Dye	Exp, msec	Binning	SNR (select points)	Average SNR
YOYO-1	2	9 x 9	192, 202, 184, 170, 181	186
TOTO-1	2	9 x 9	69, 84, 72	75

FIG. 3

FIG. 4B

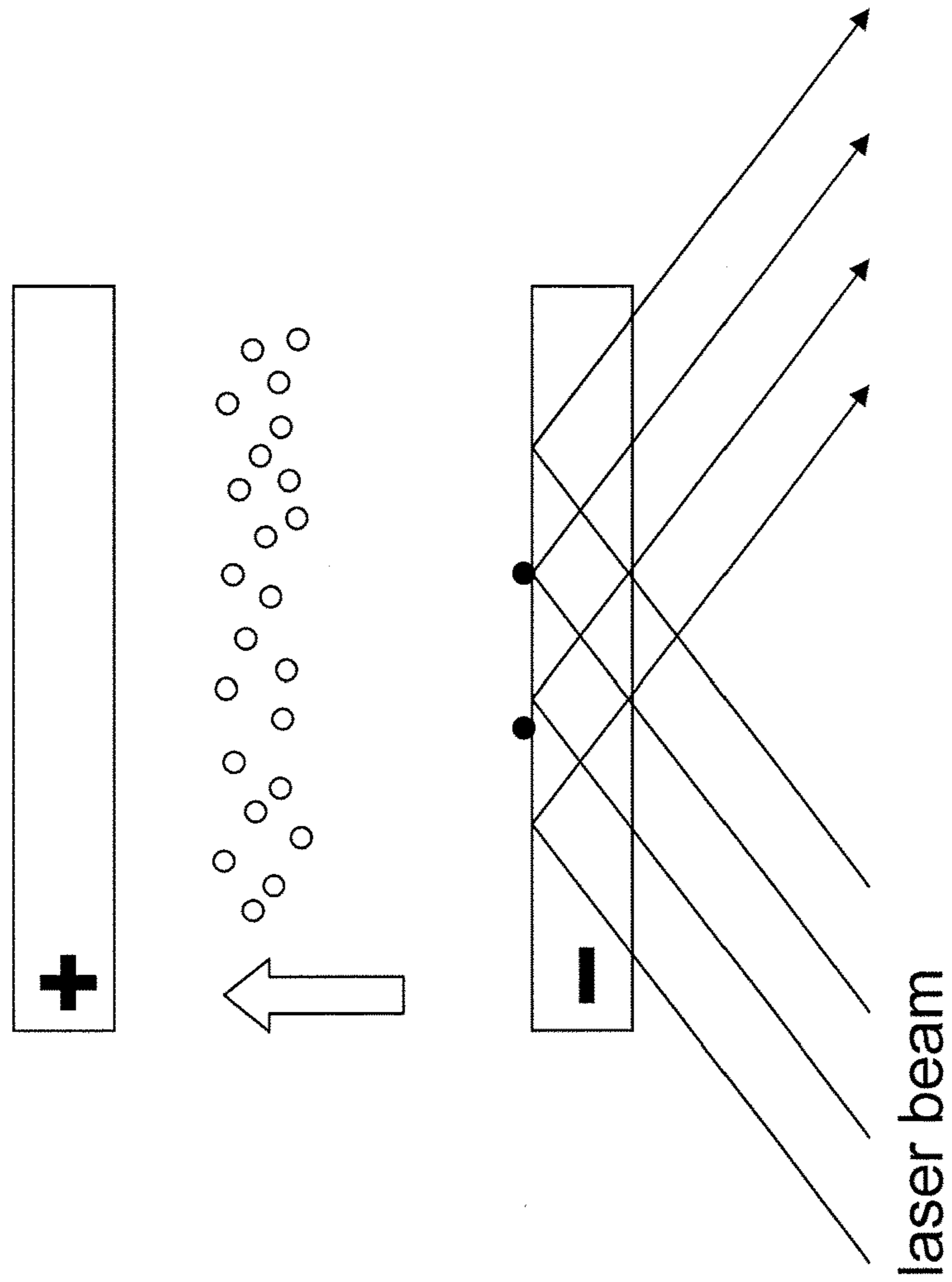
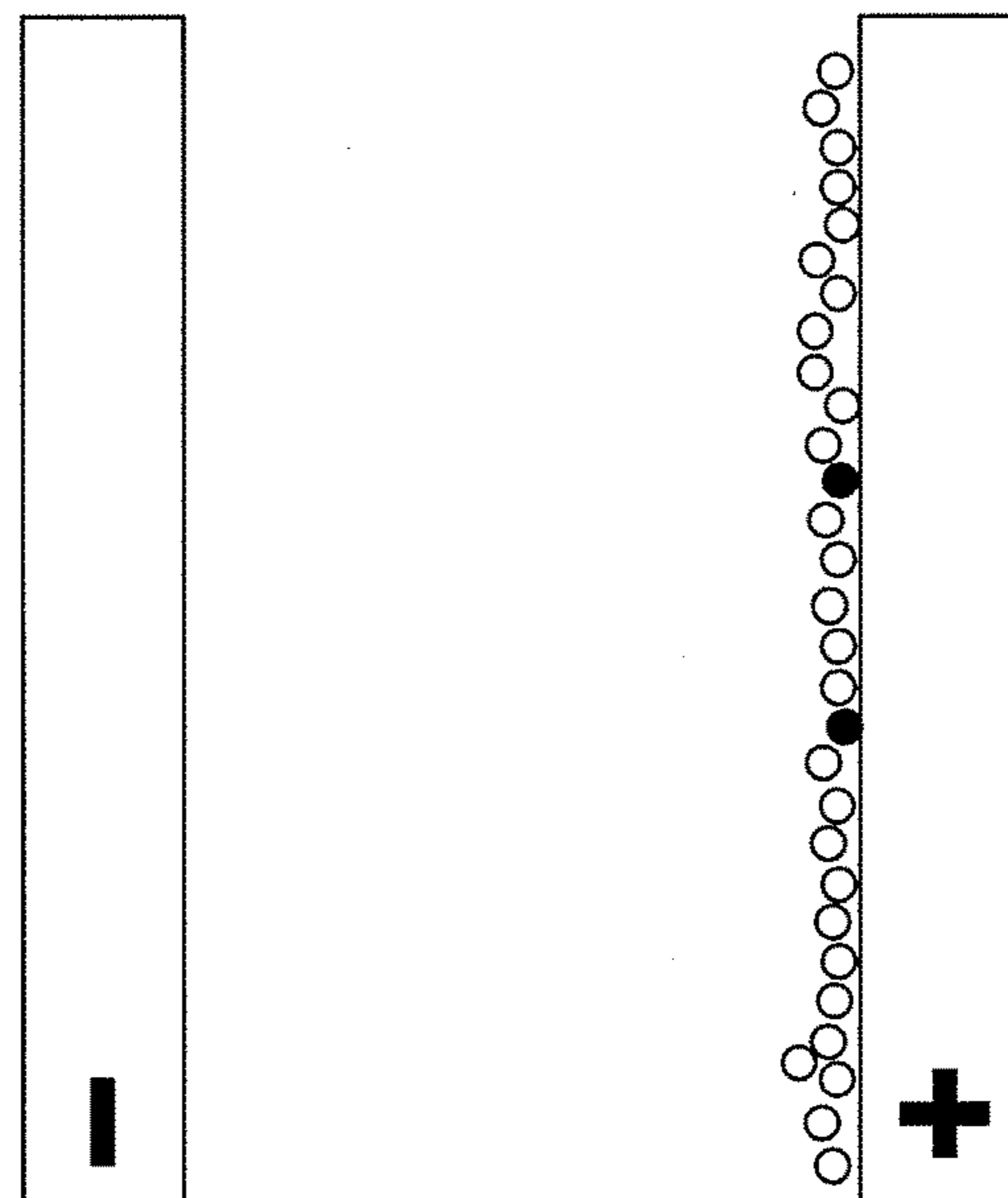


FIG. 4A



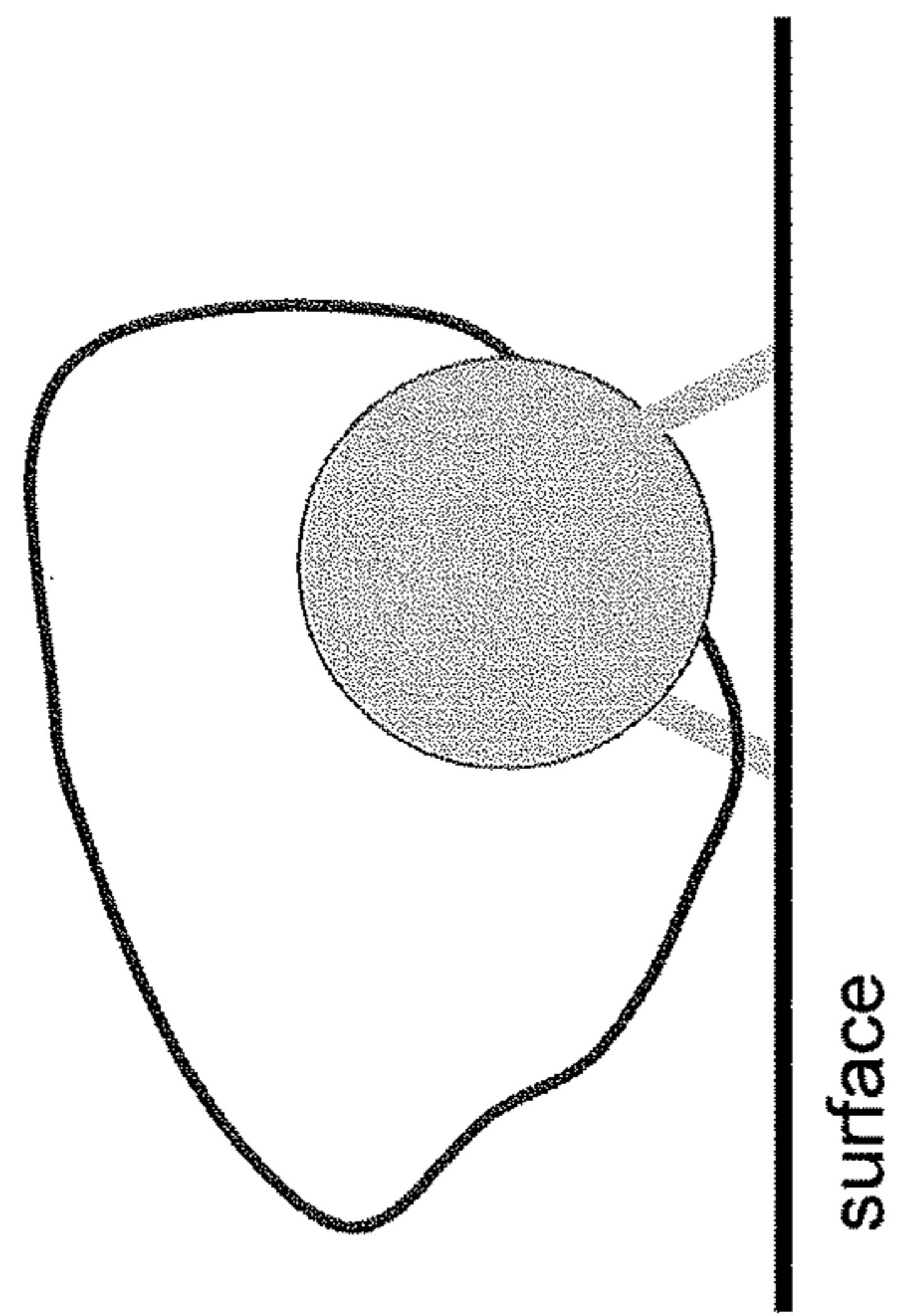


FIG. 5A

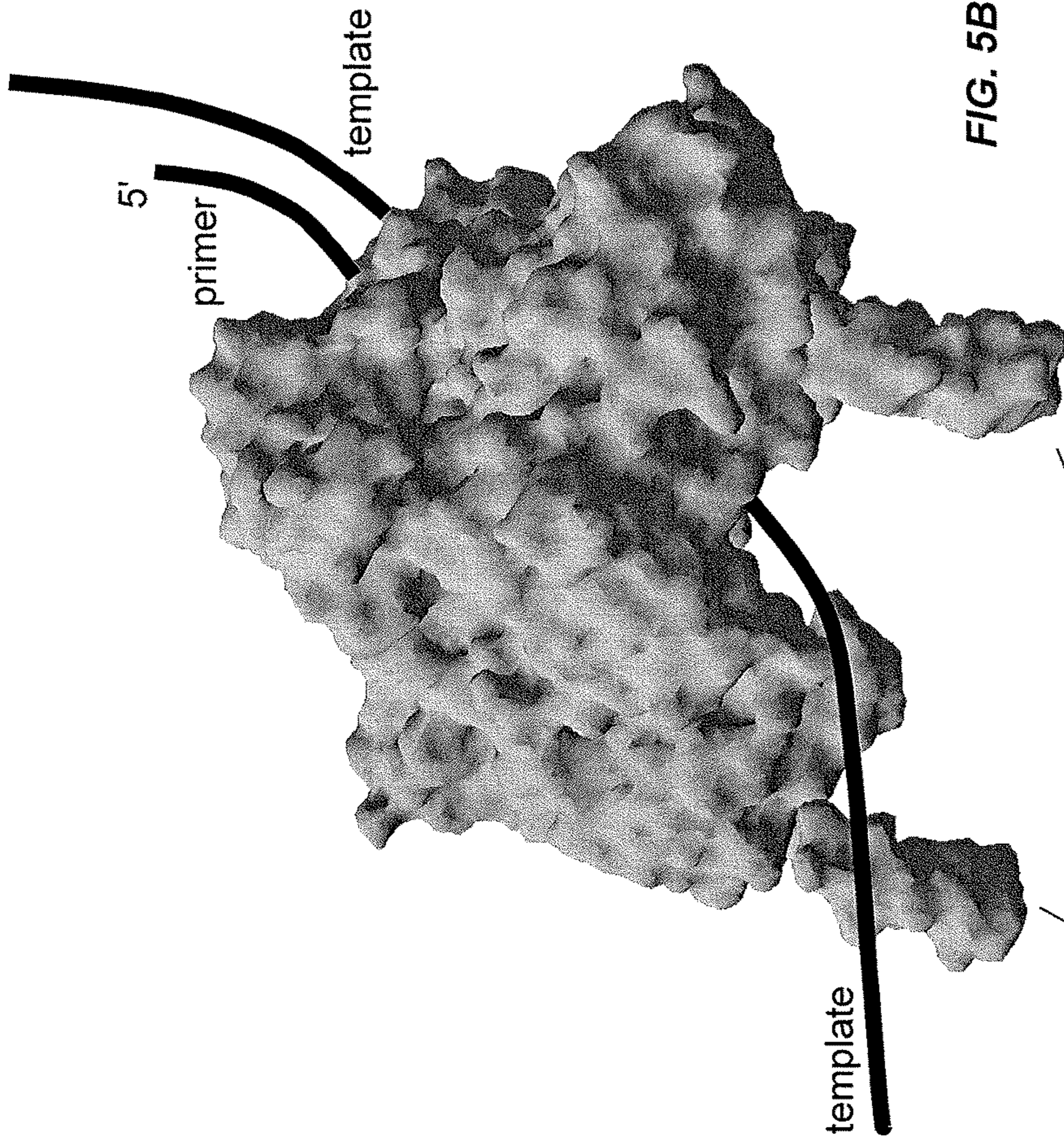


FIG. 5B

bio  
Biotin loop  
(SSGGLNDFEFAQKIEWHEDTGSS)

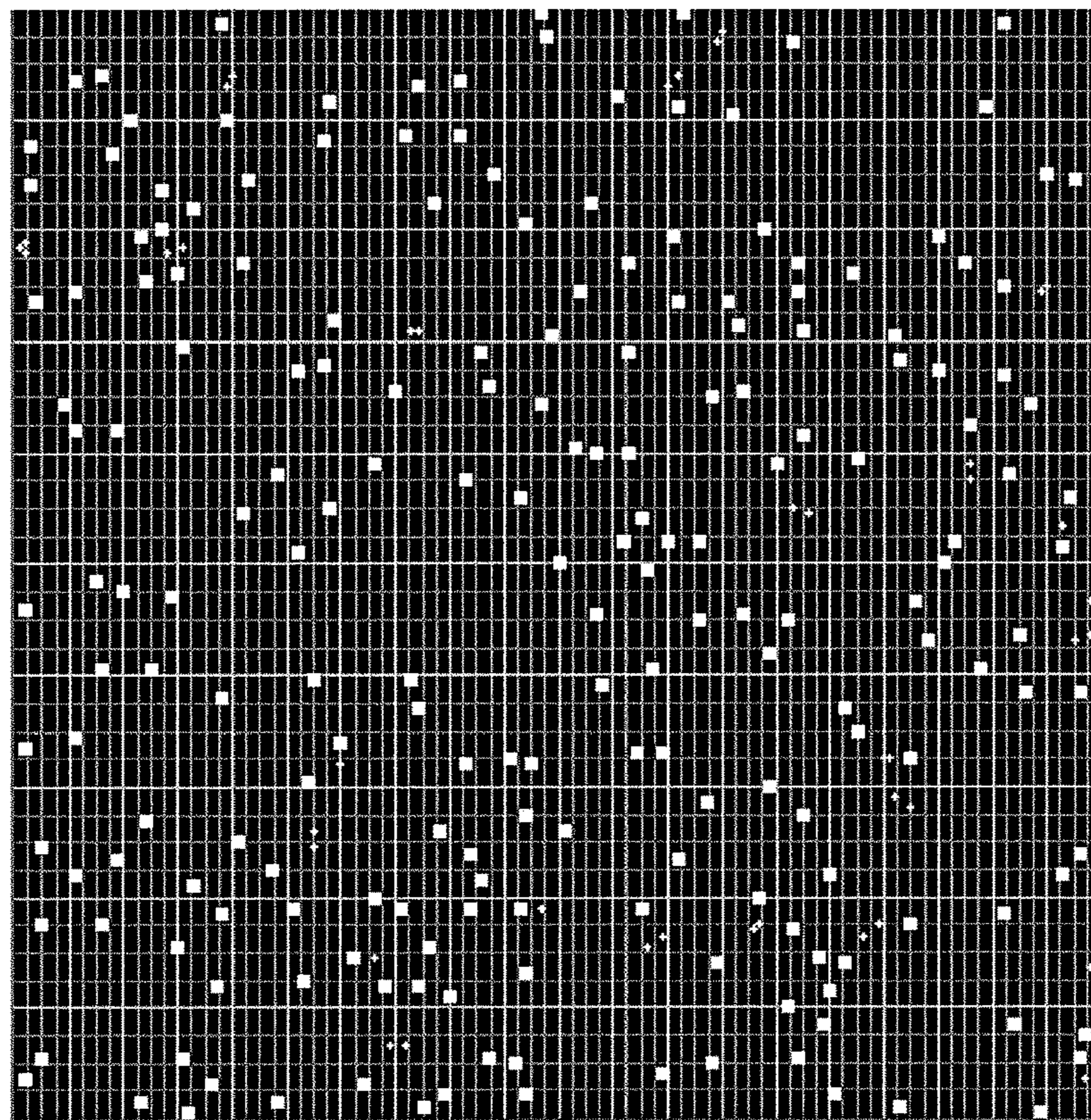


FIG. 6A

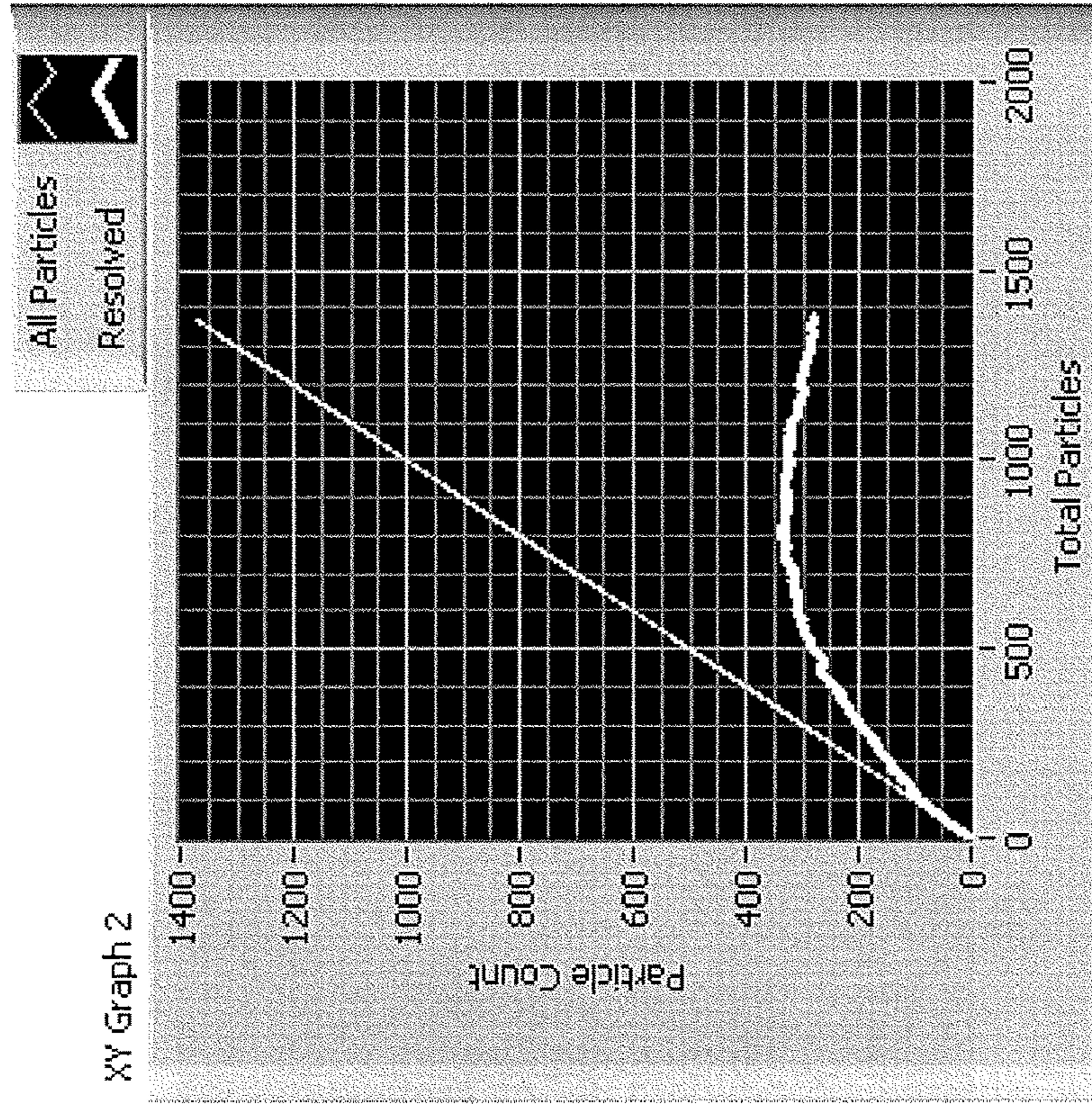


FIG. 6B



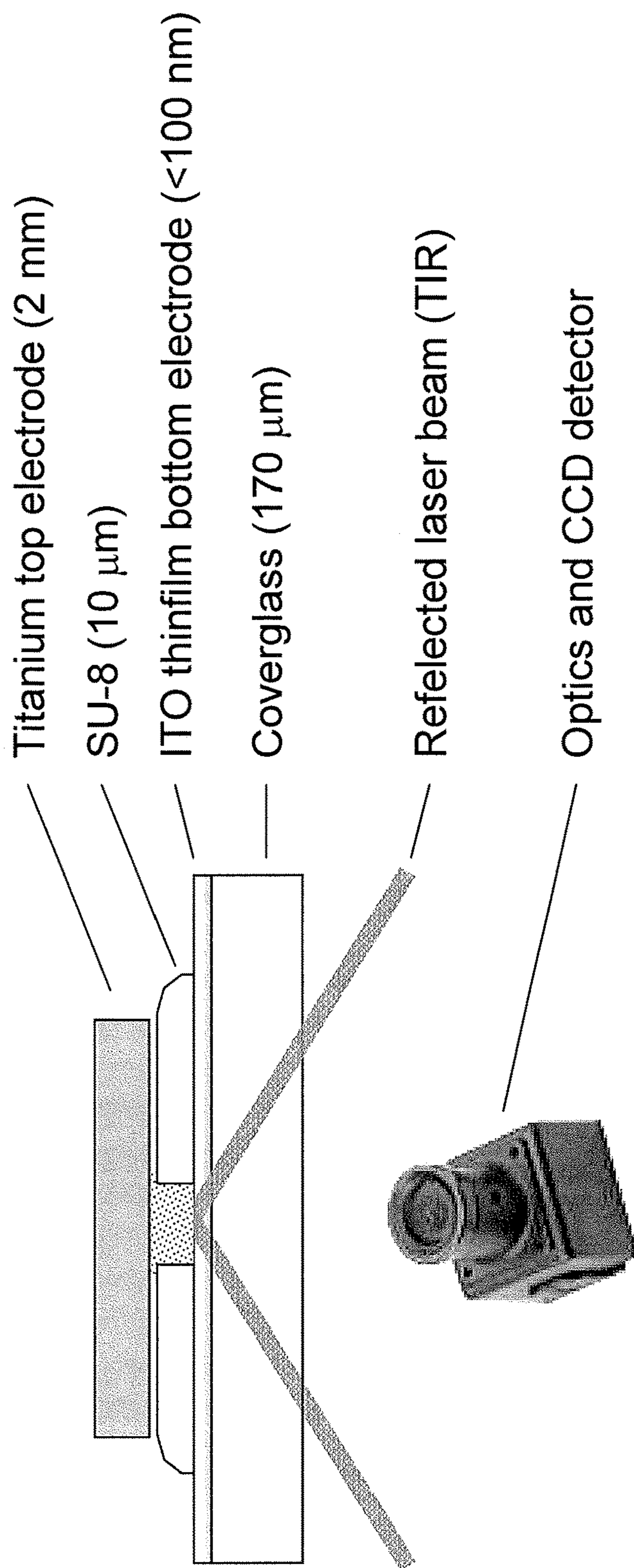


FIG. 7

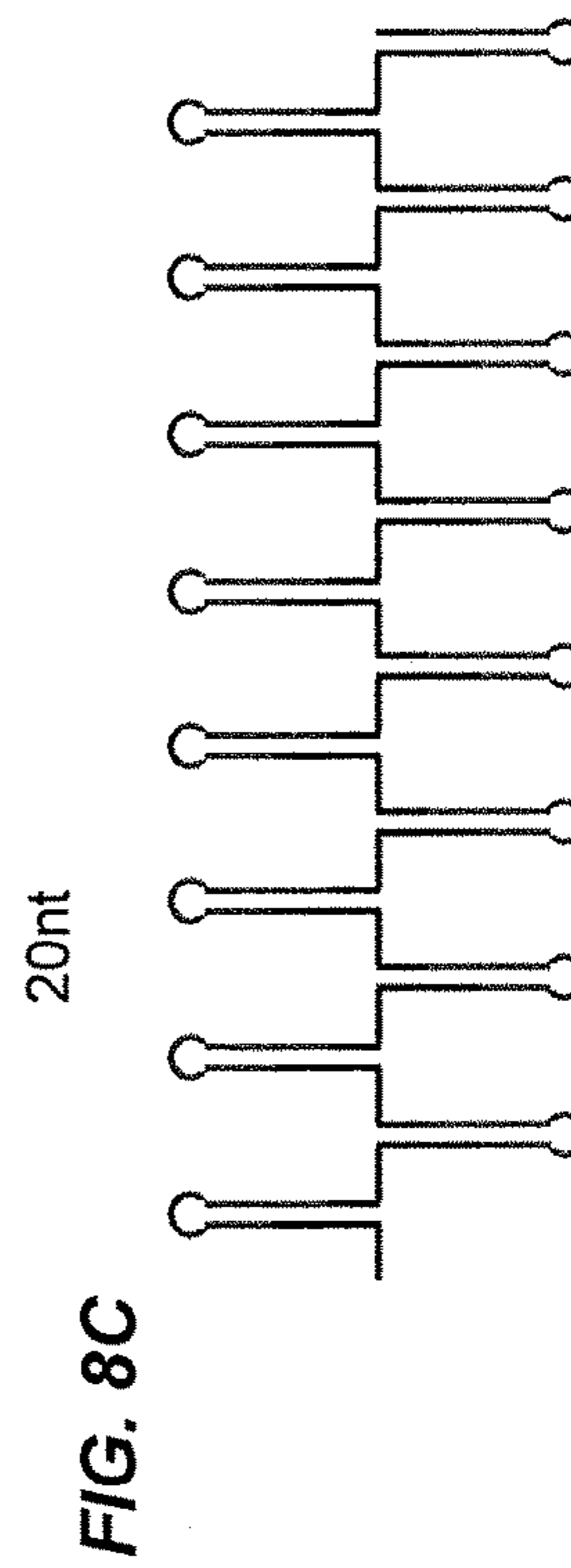
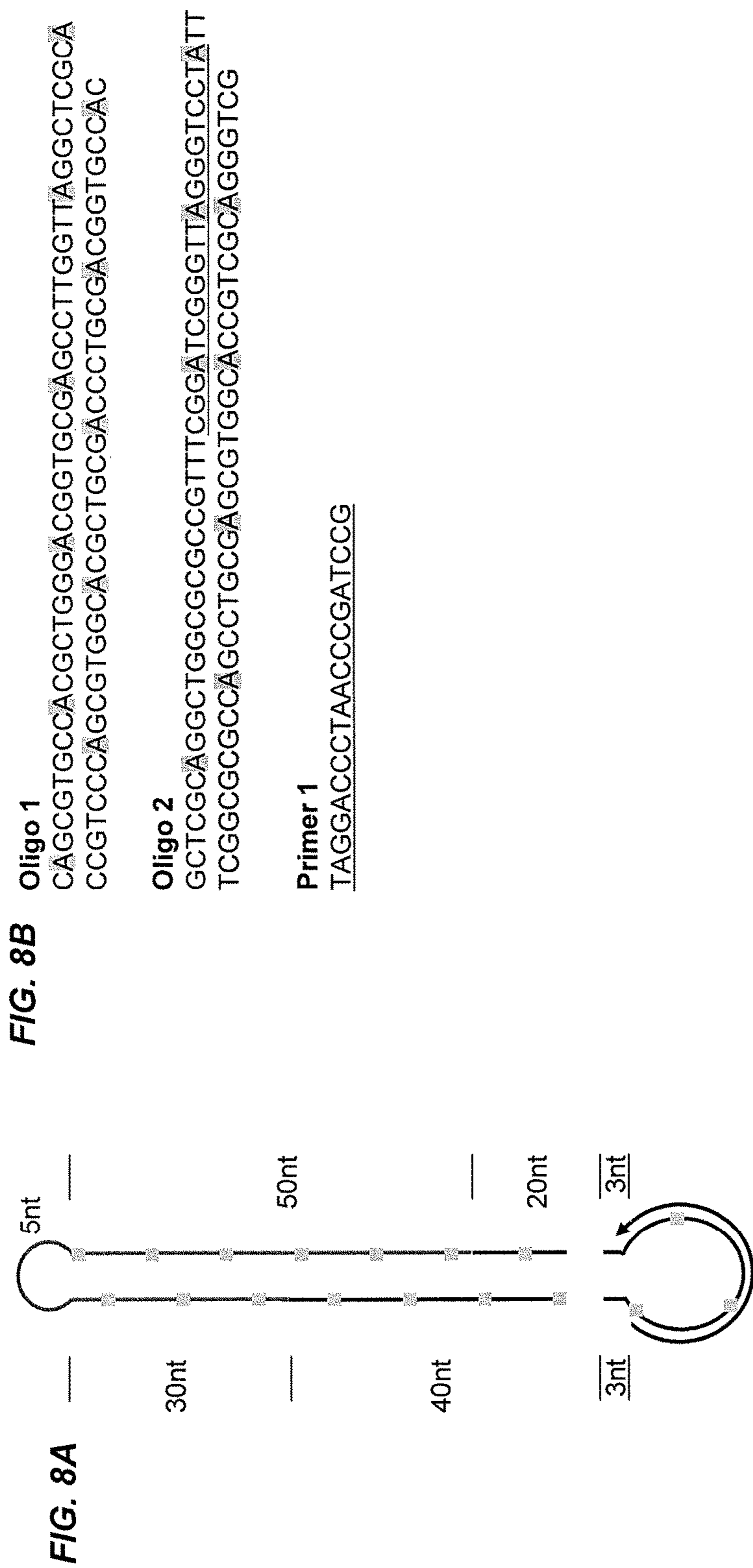




FIG. 9B

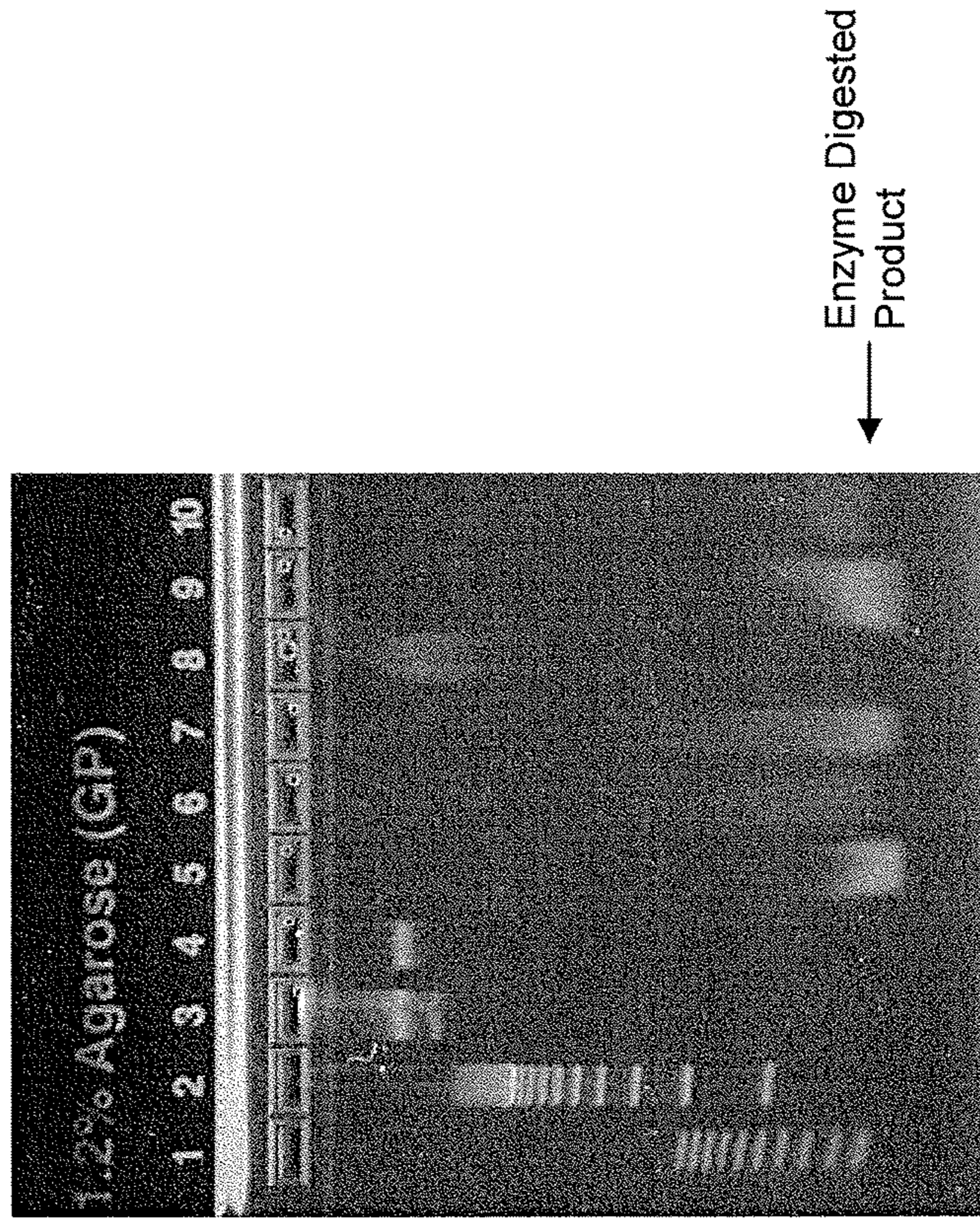
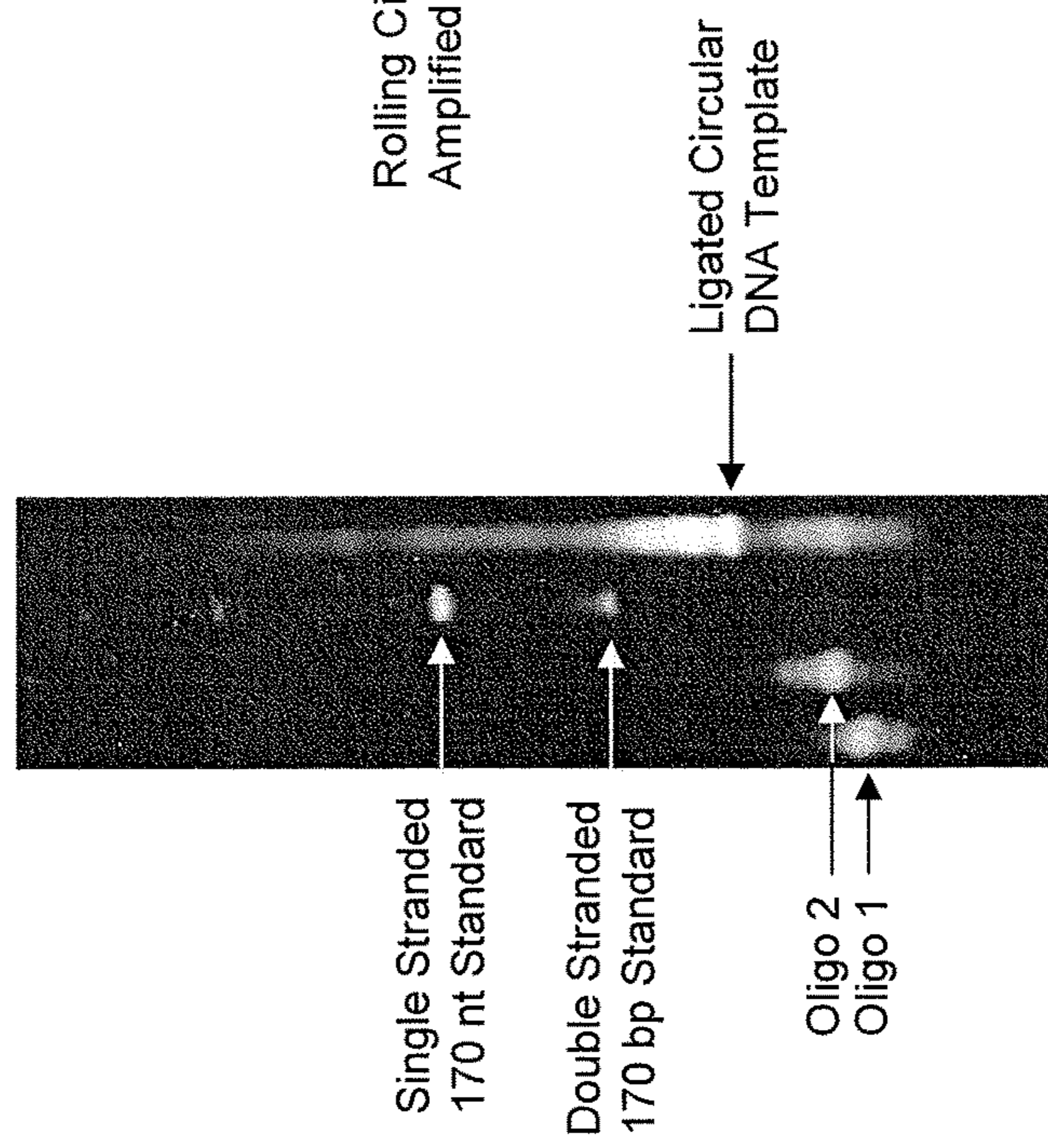
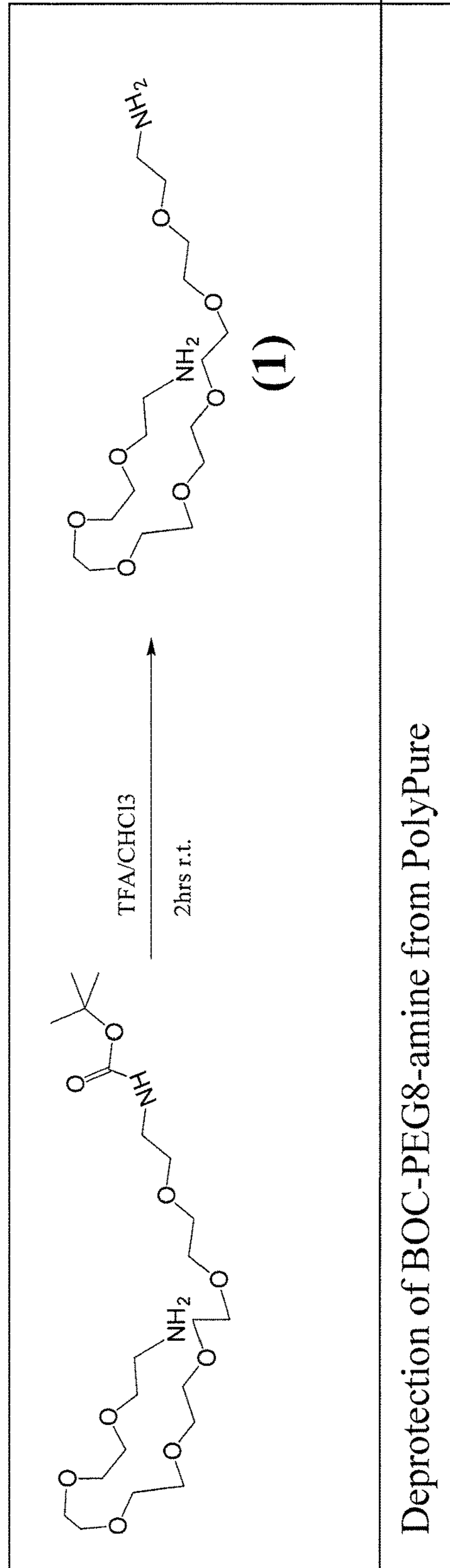


FIG. 9A





Deprotection of BOC-PEG8-amine from PolyPure

FIG. 10

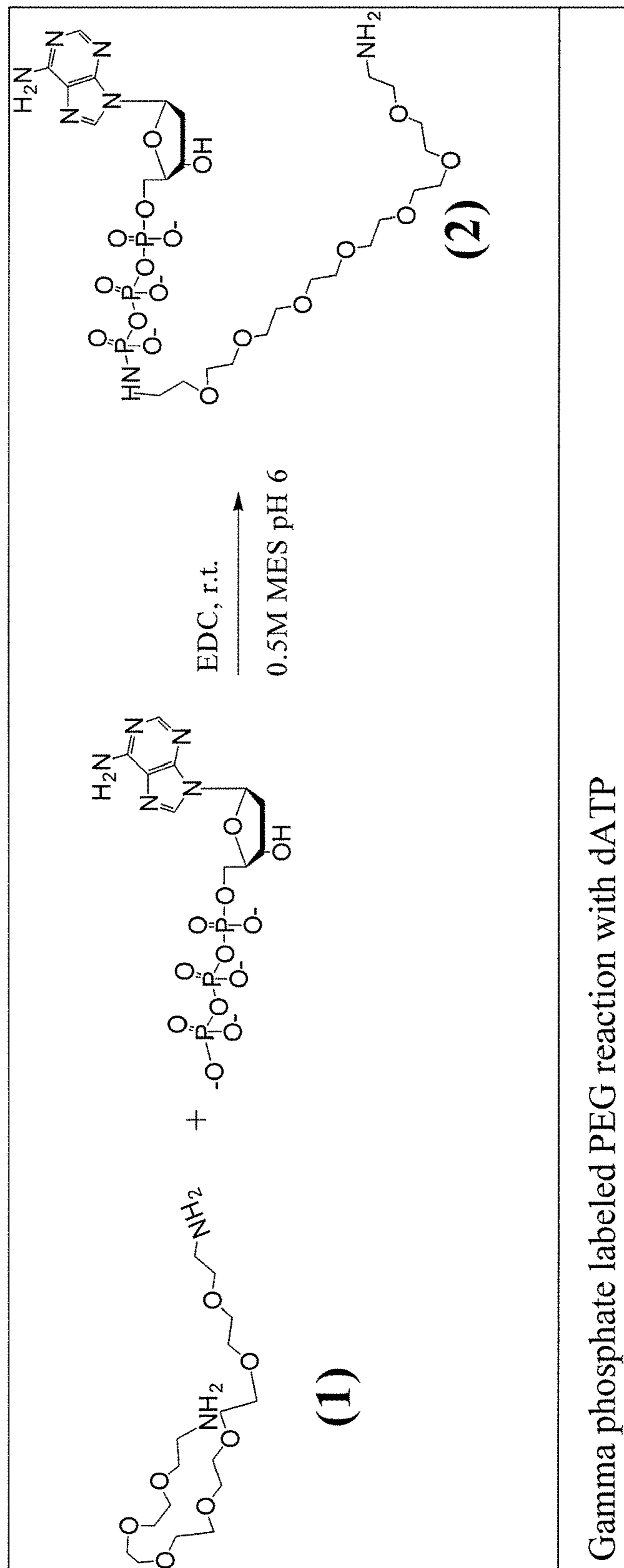


FIG. 11

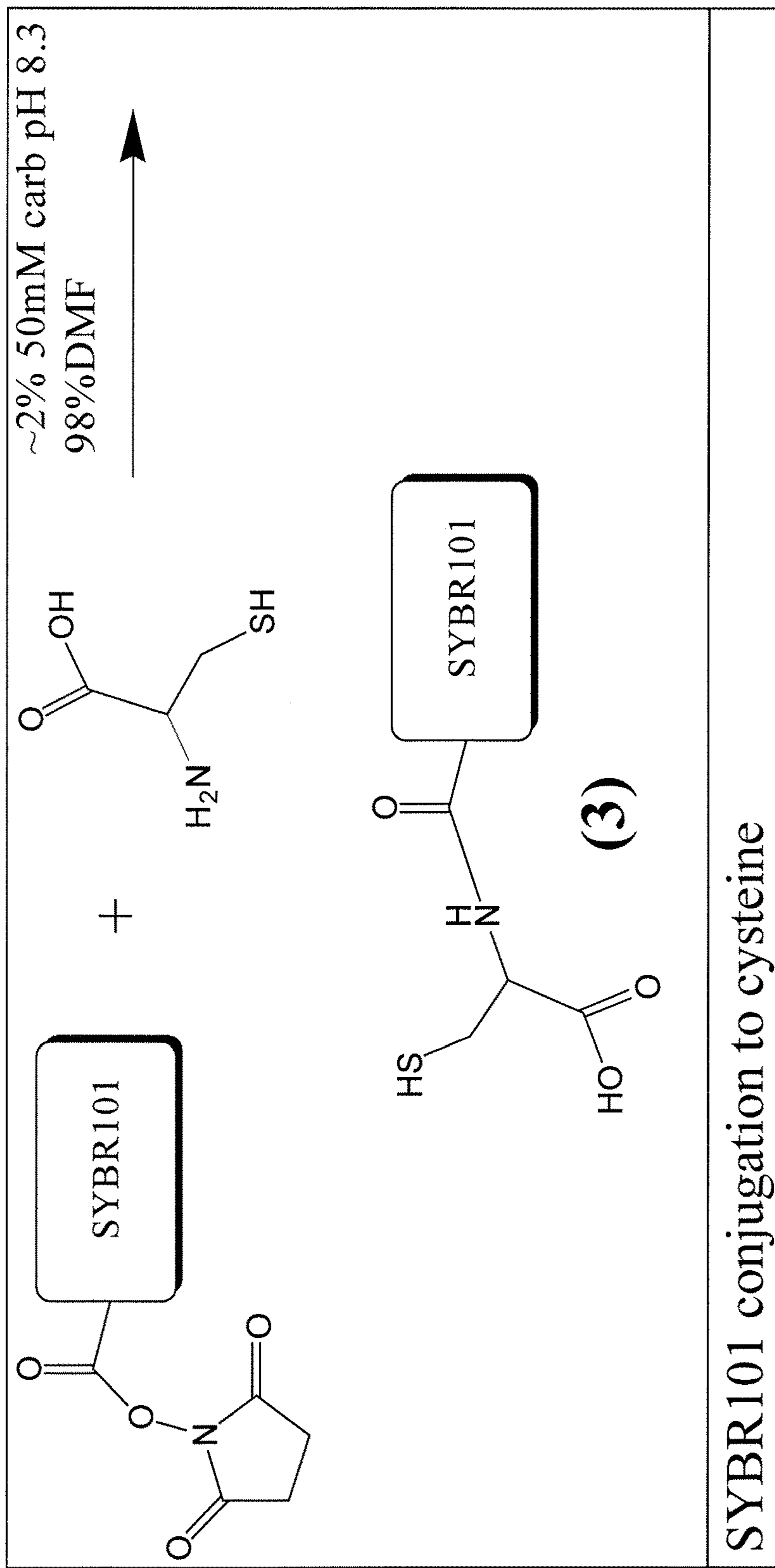


FIG. 12

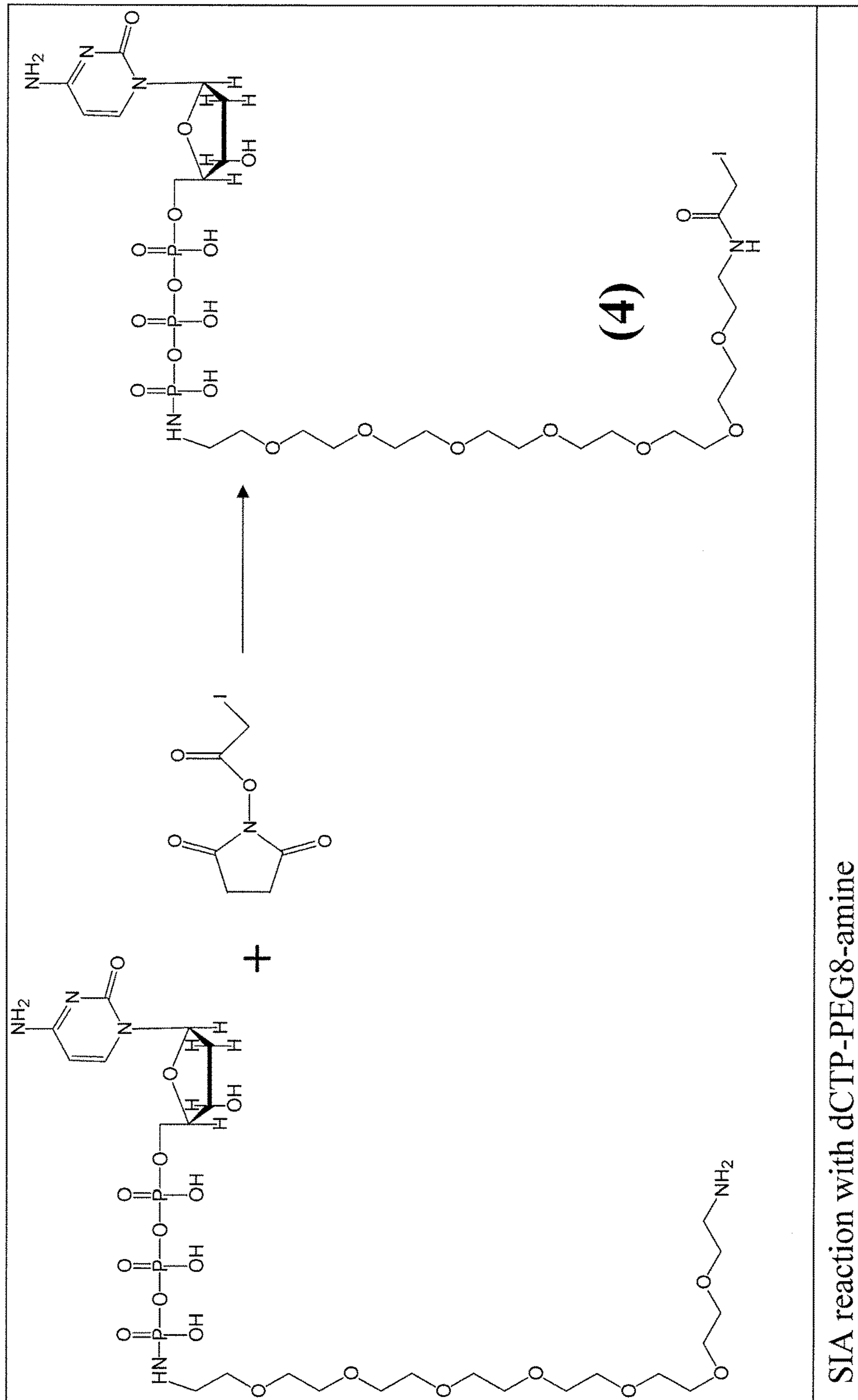
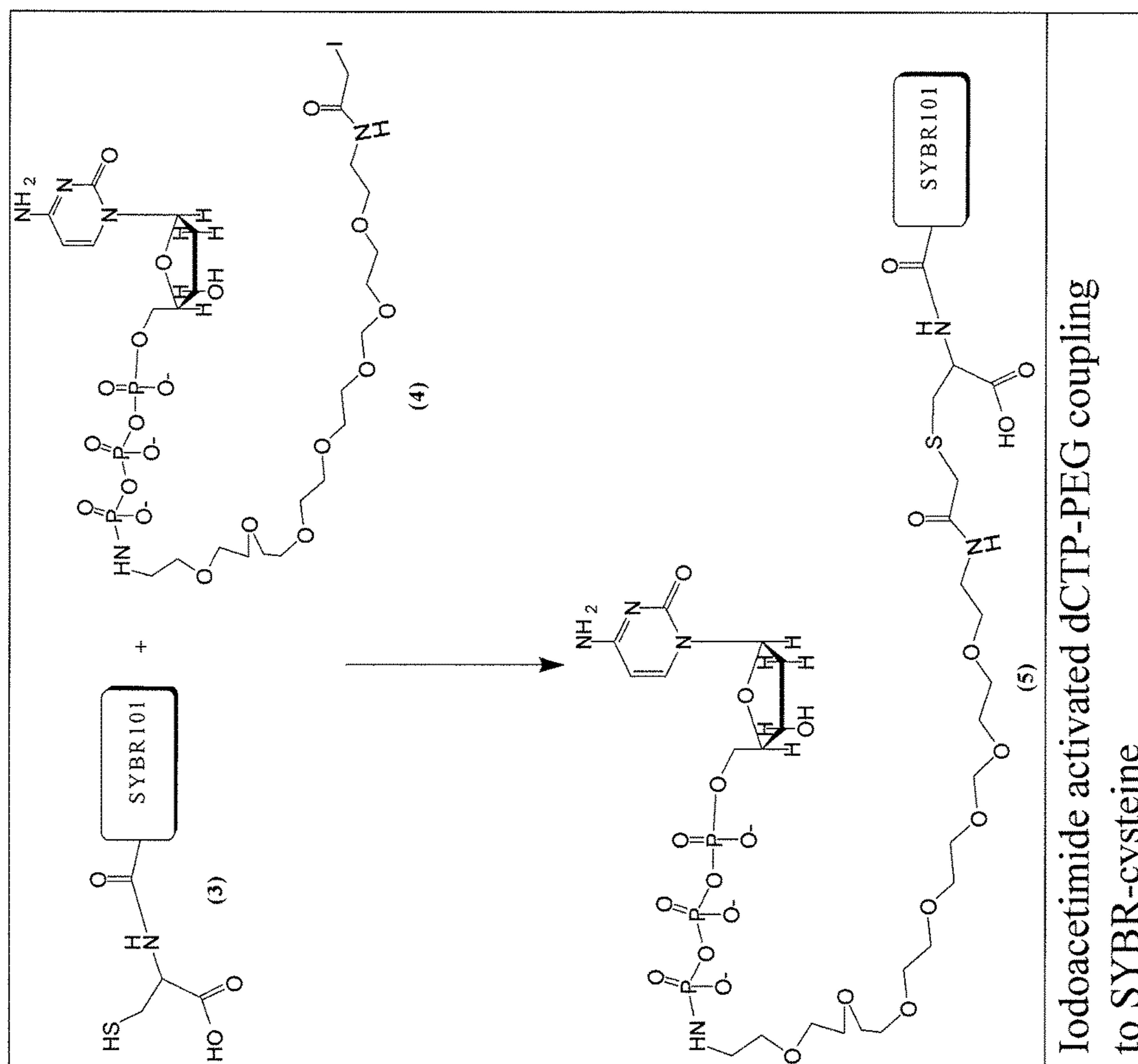


FIG. 13





Iodoacetamide activated dCTP-PEG coupling to SYBR-cysteine

FIG. 14

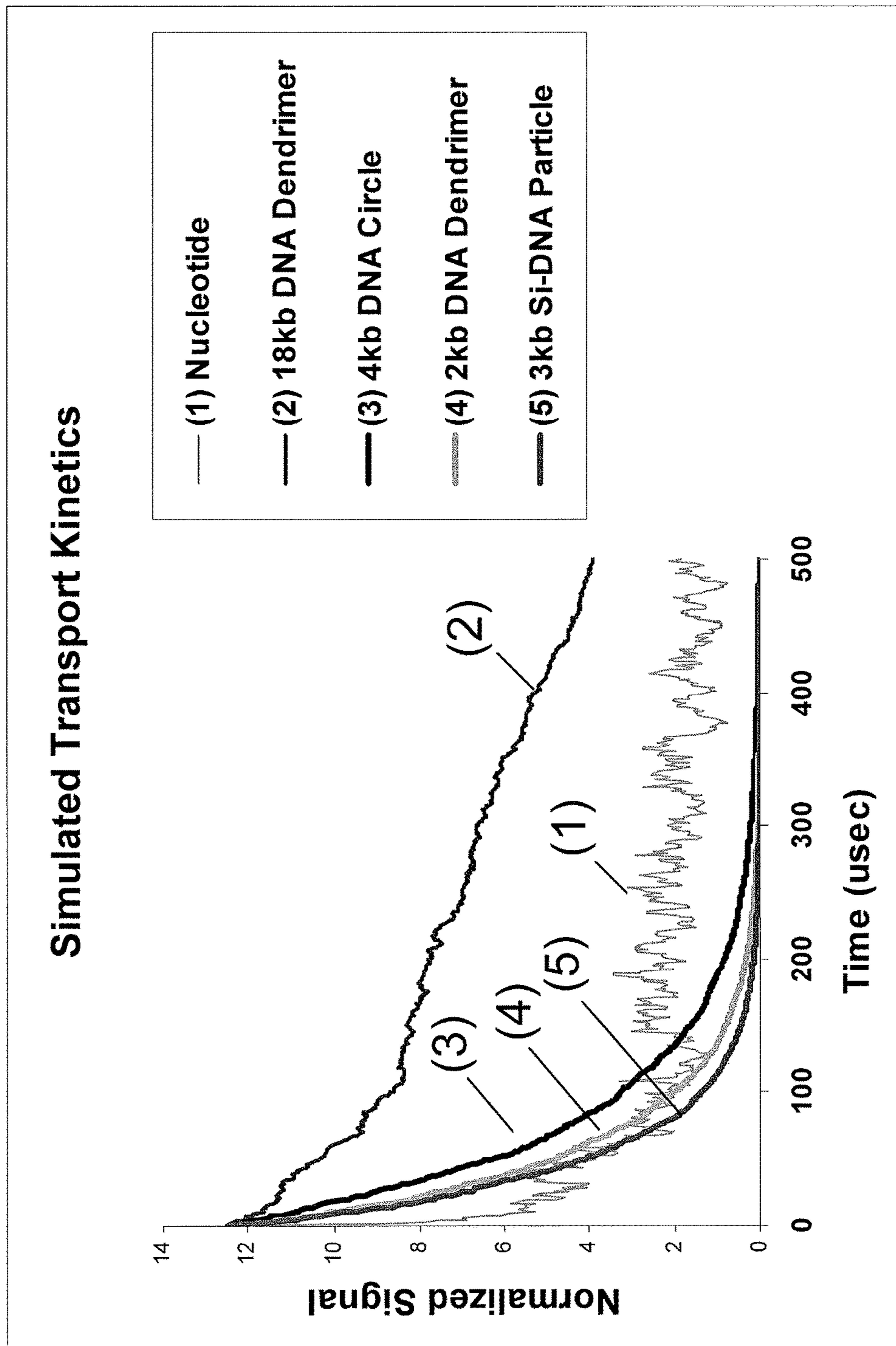
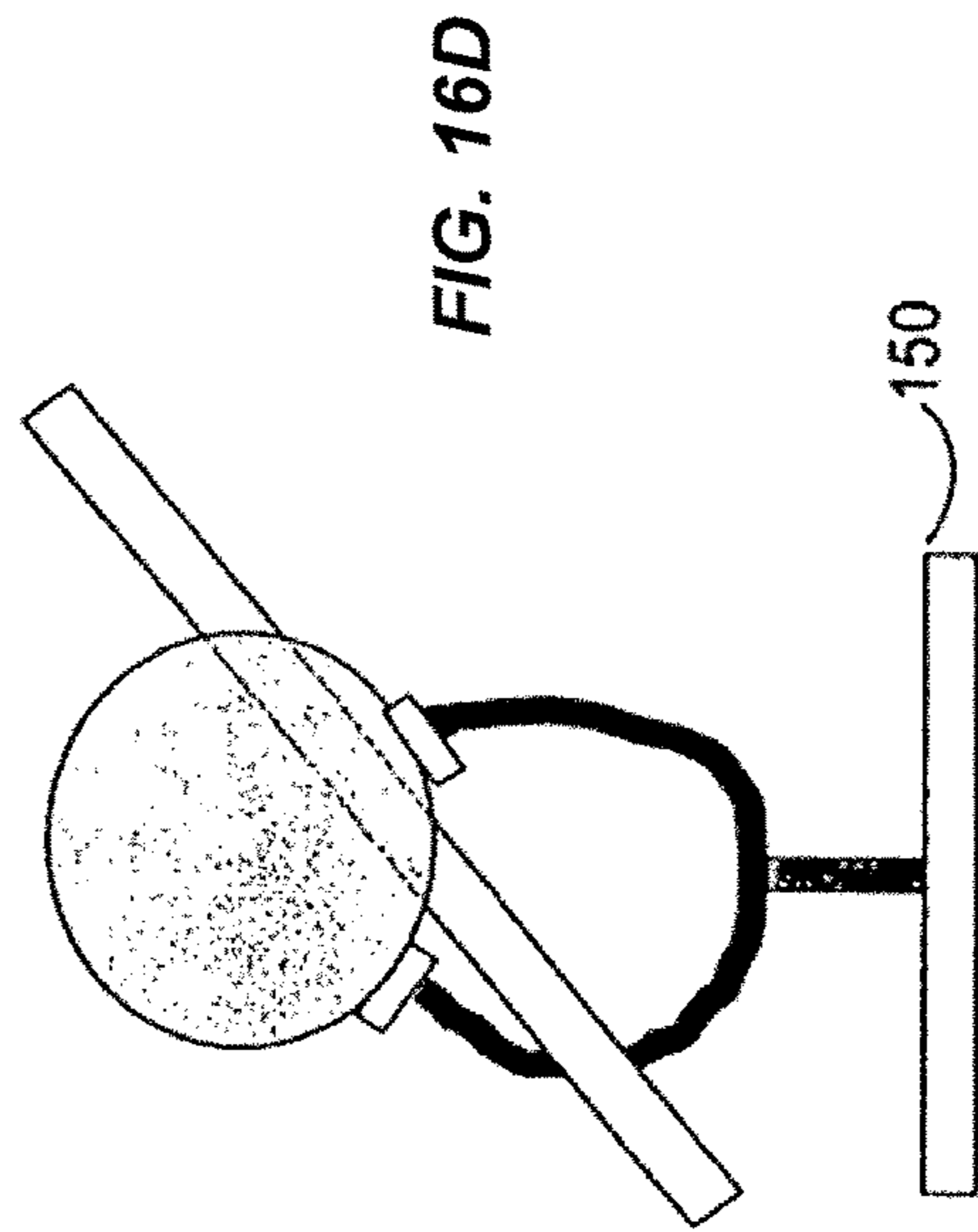
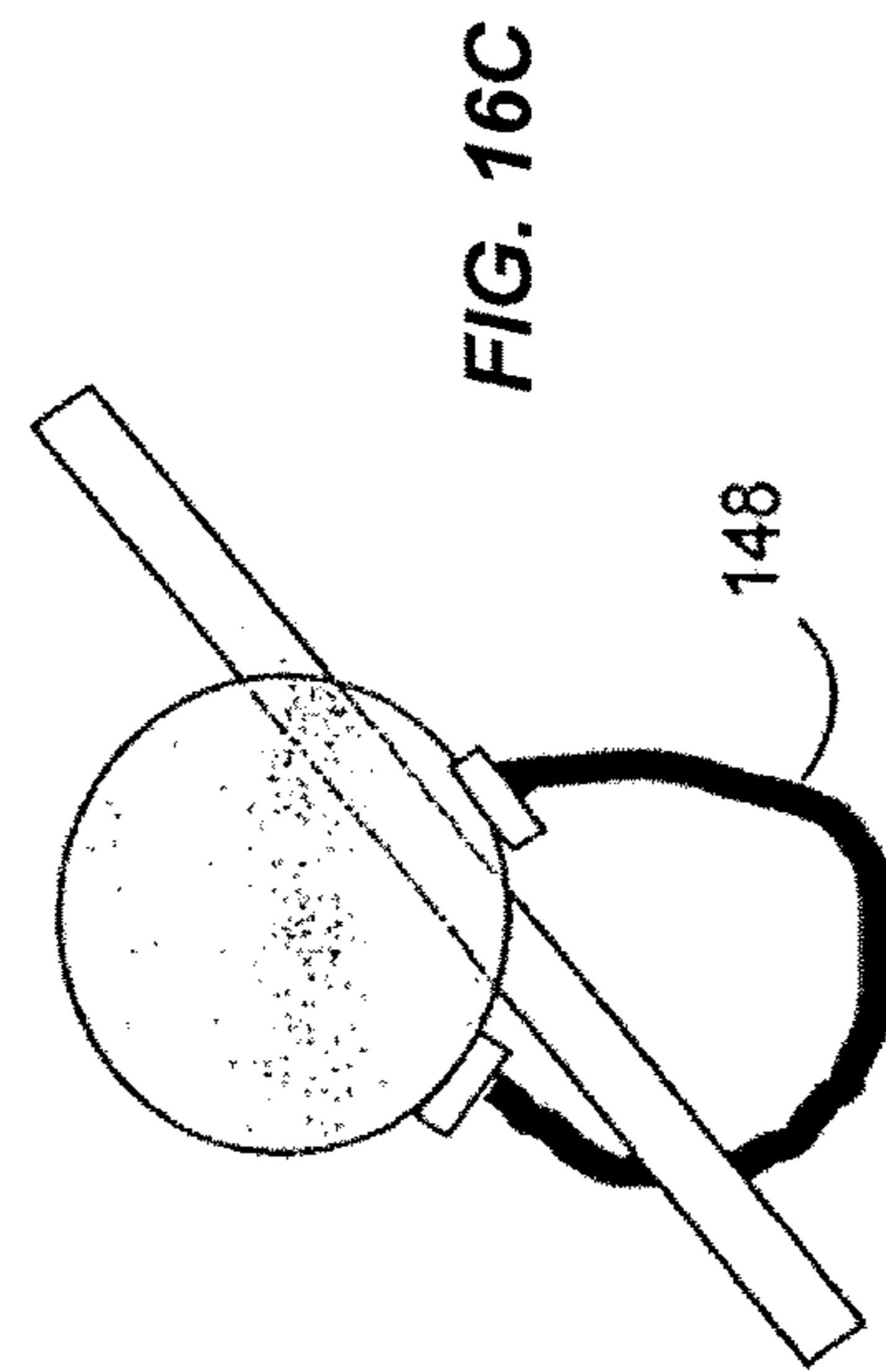
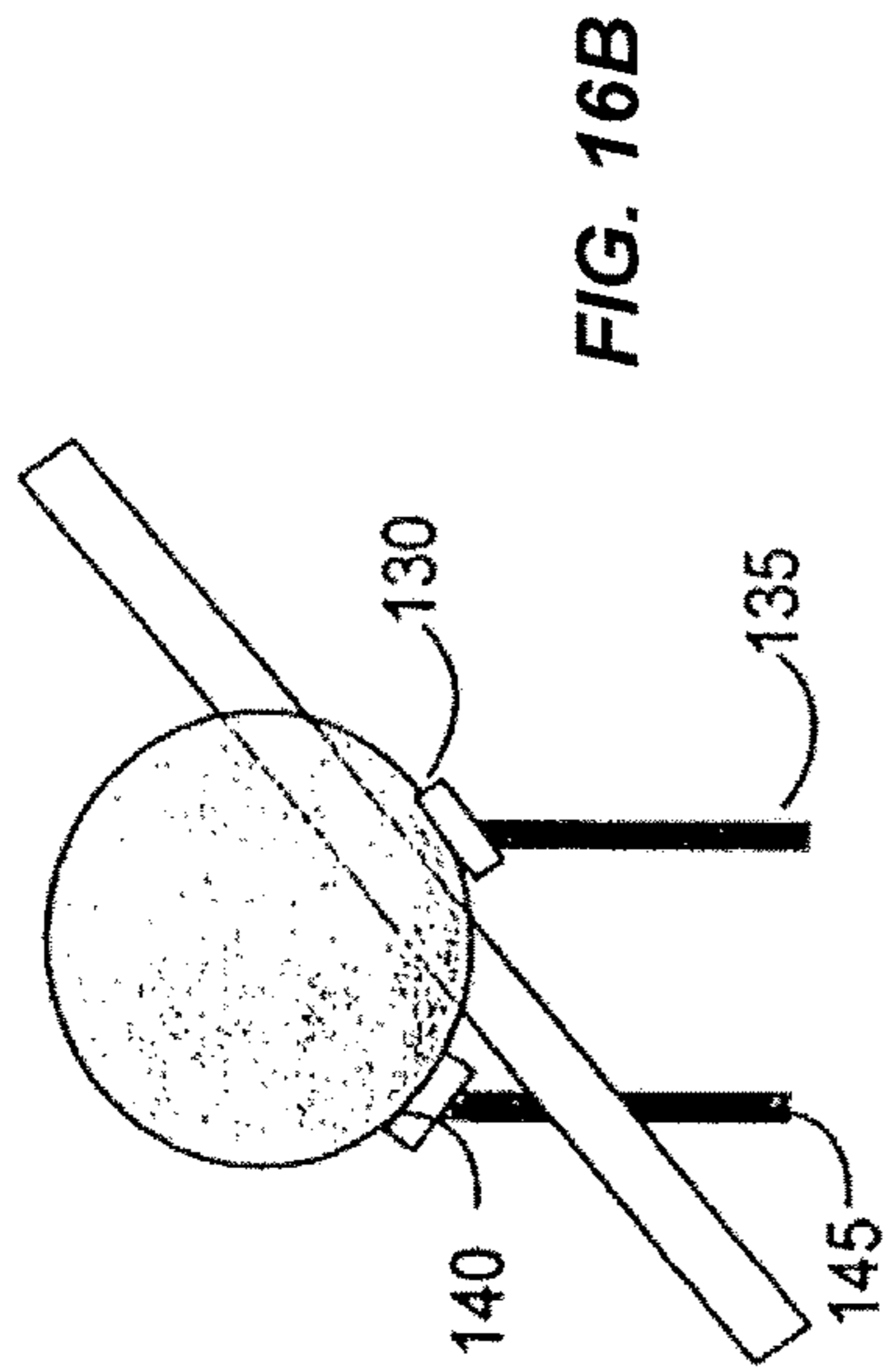
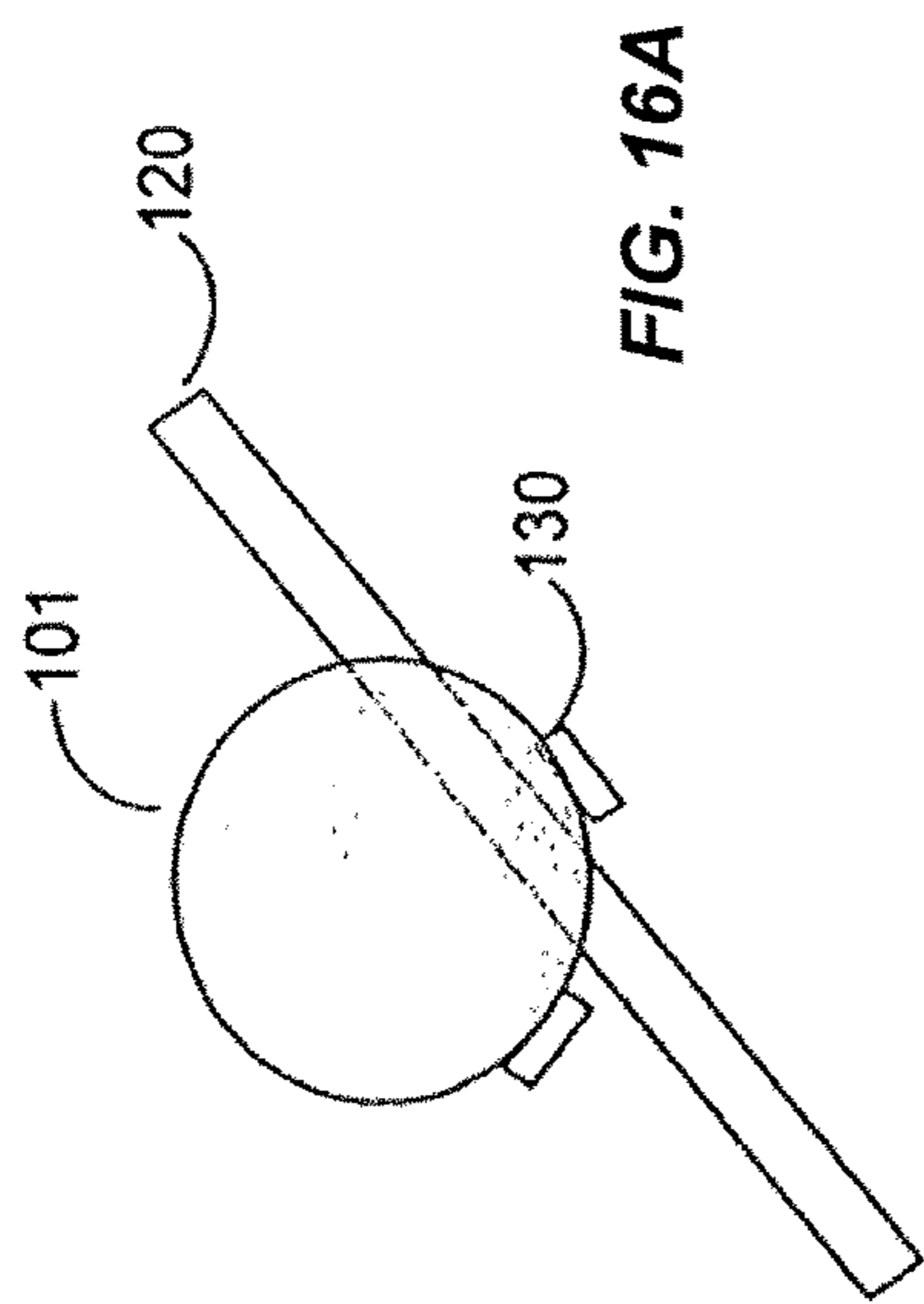


FIG. 15



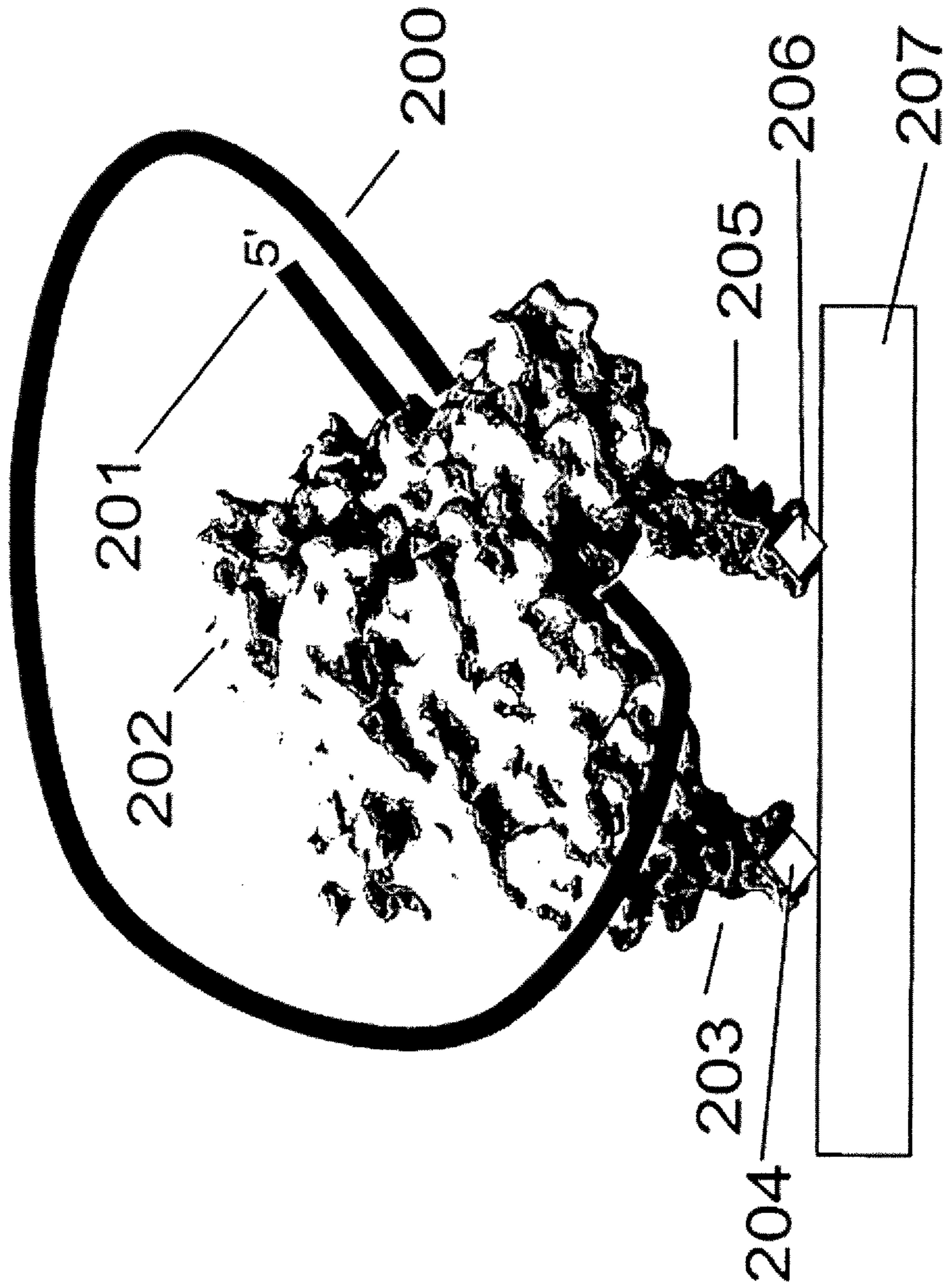
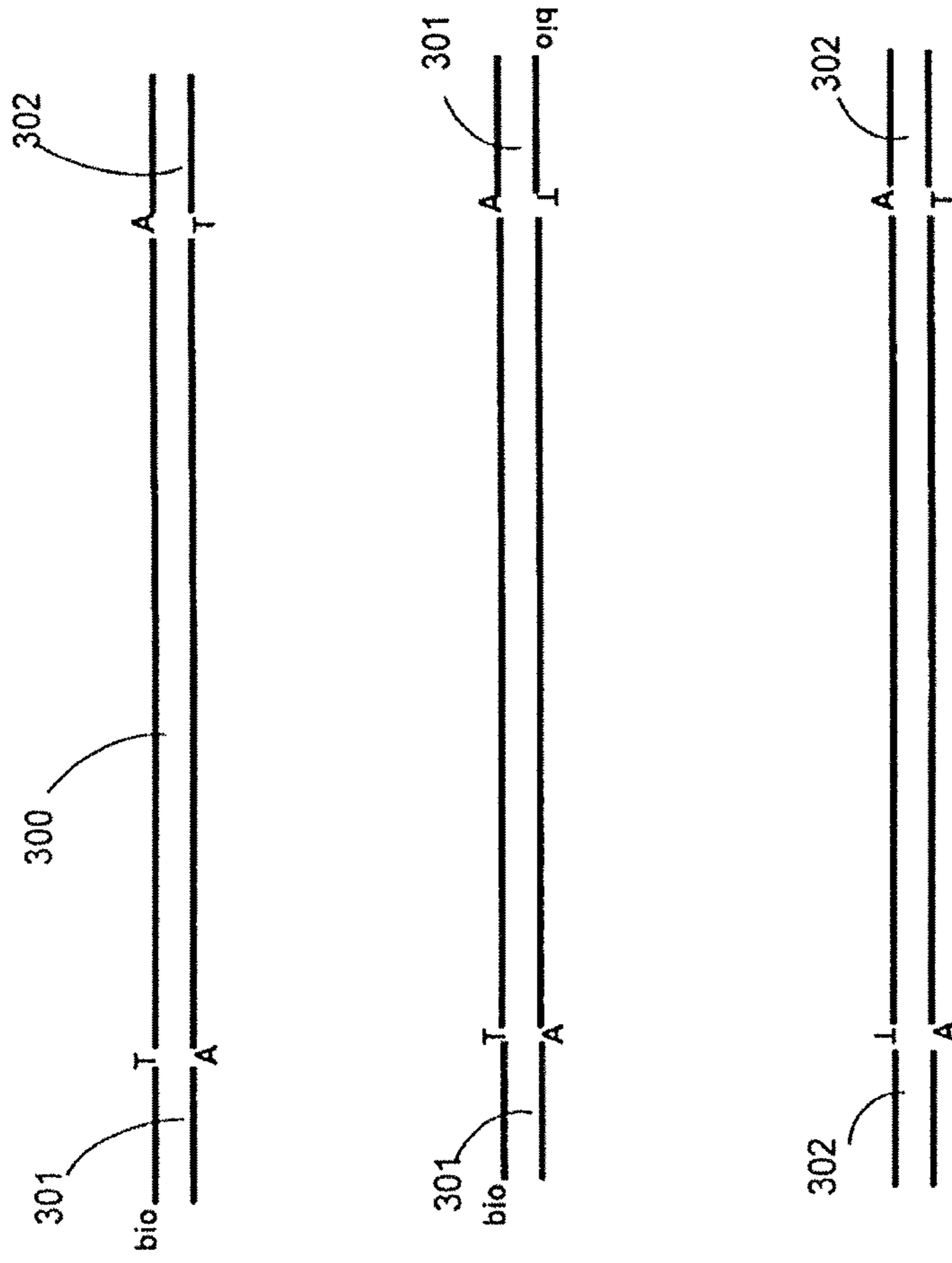


FIG. 17

# DNA Sample Prep

## adaptor ligation



## PURIFY 20KB FRAGMENTS

FIG. 18

# DNA Sample Prep

strand isolation

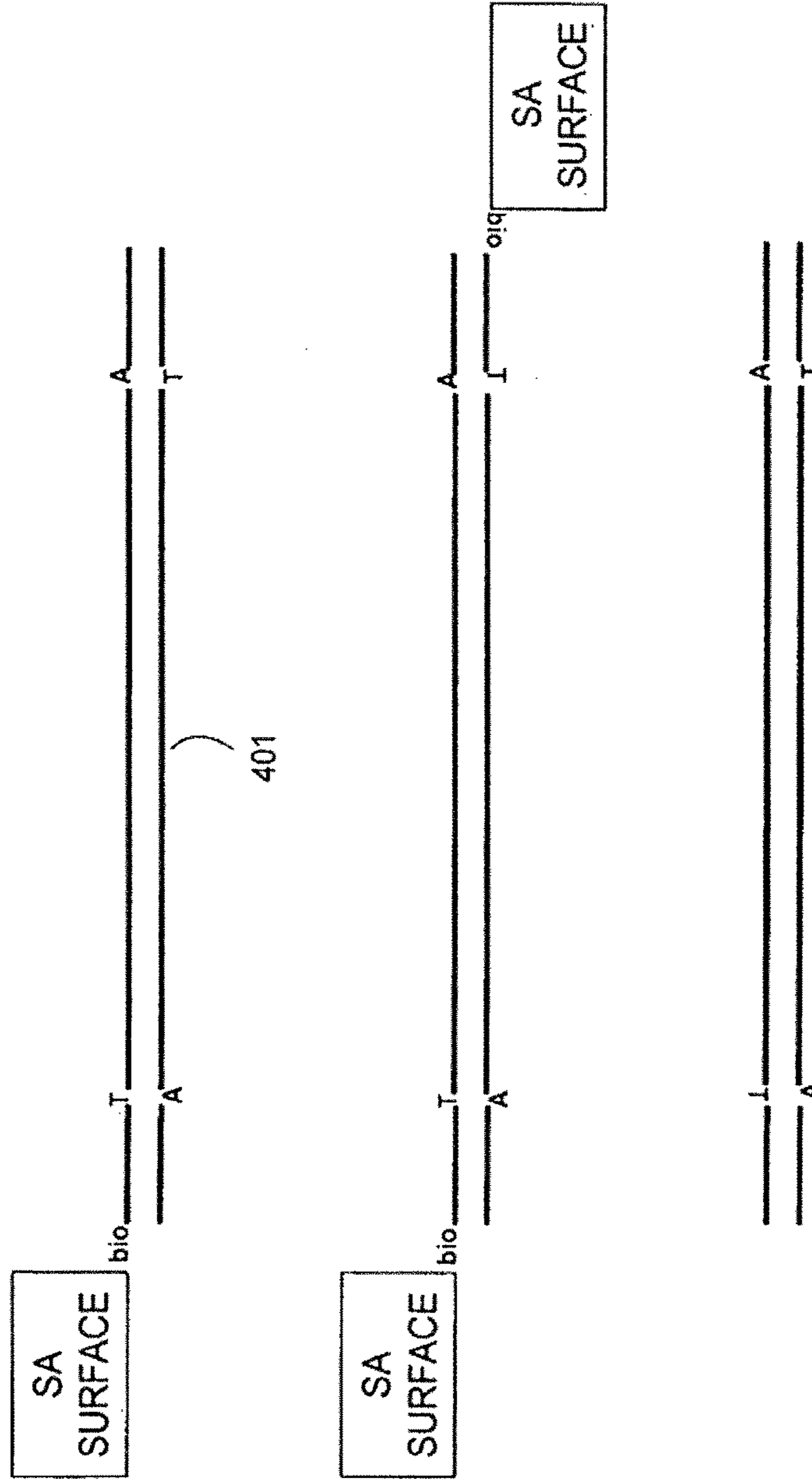


FIG. 19

# DNA Sample Prep circularization

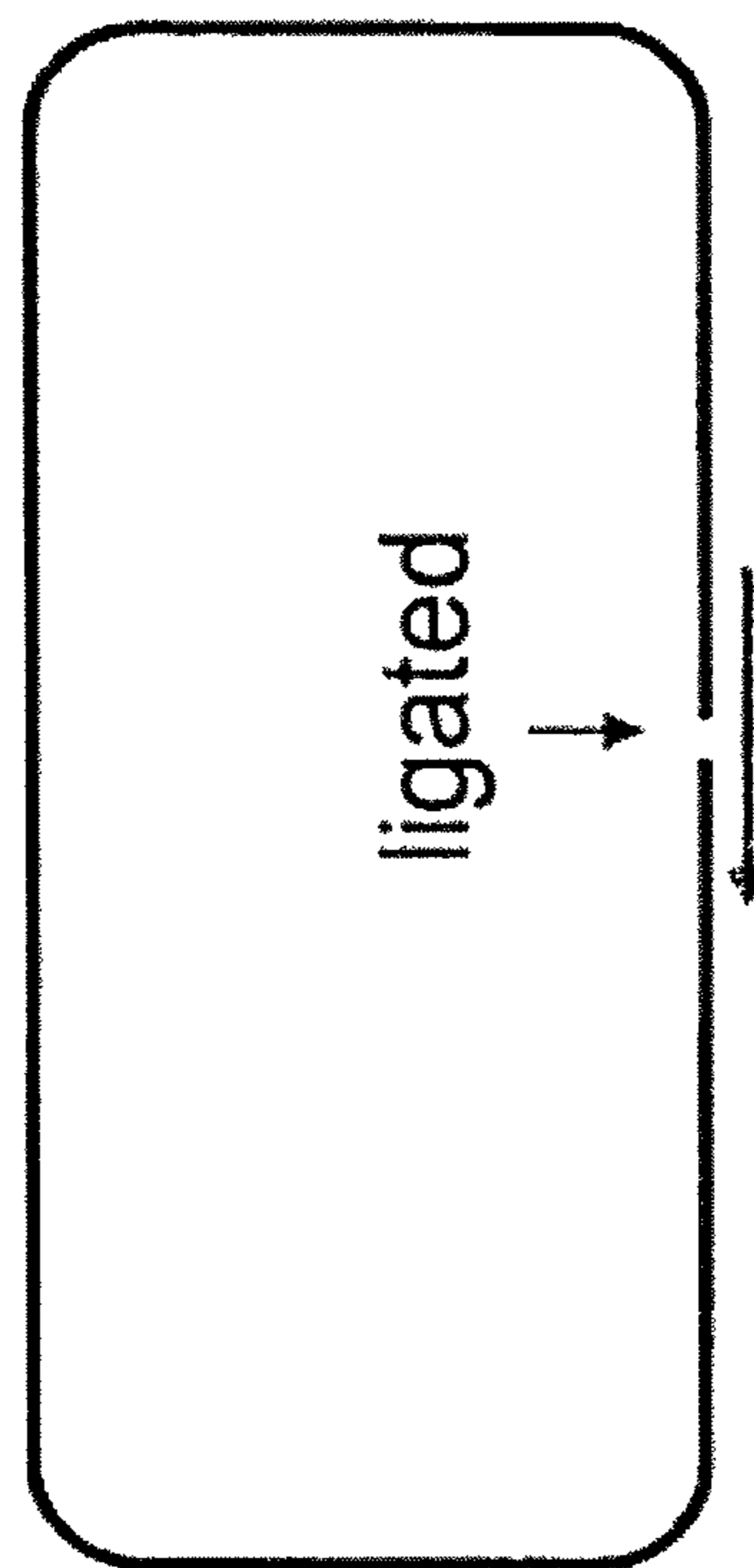


FIG. 20

# Single Molecule Configuration DNA-protein complexes

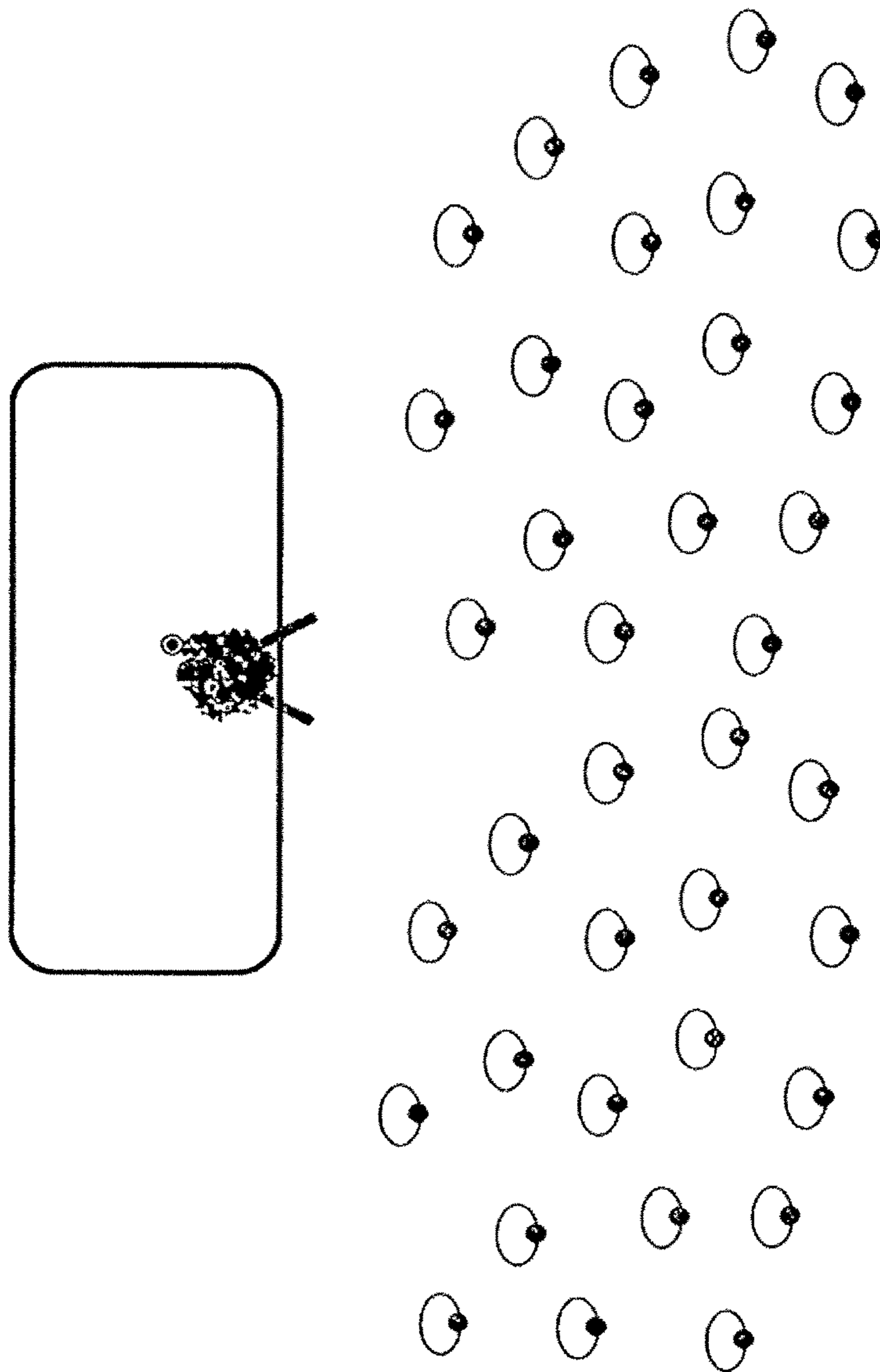


FIG. 21



# Single-Molecule Isolation

bind to coverslip

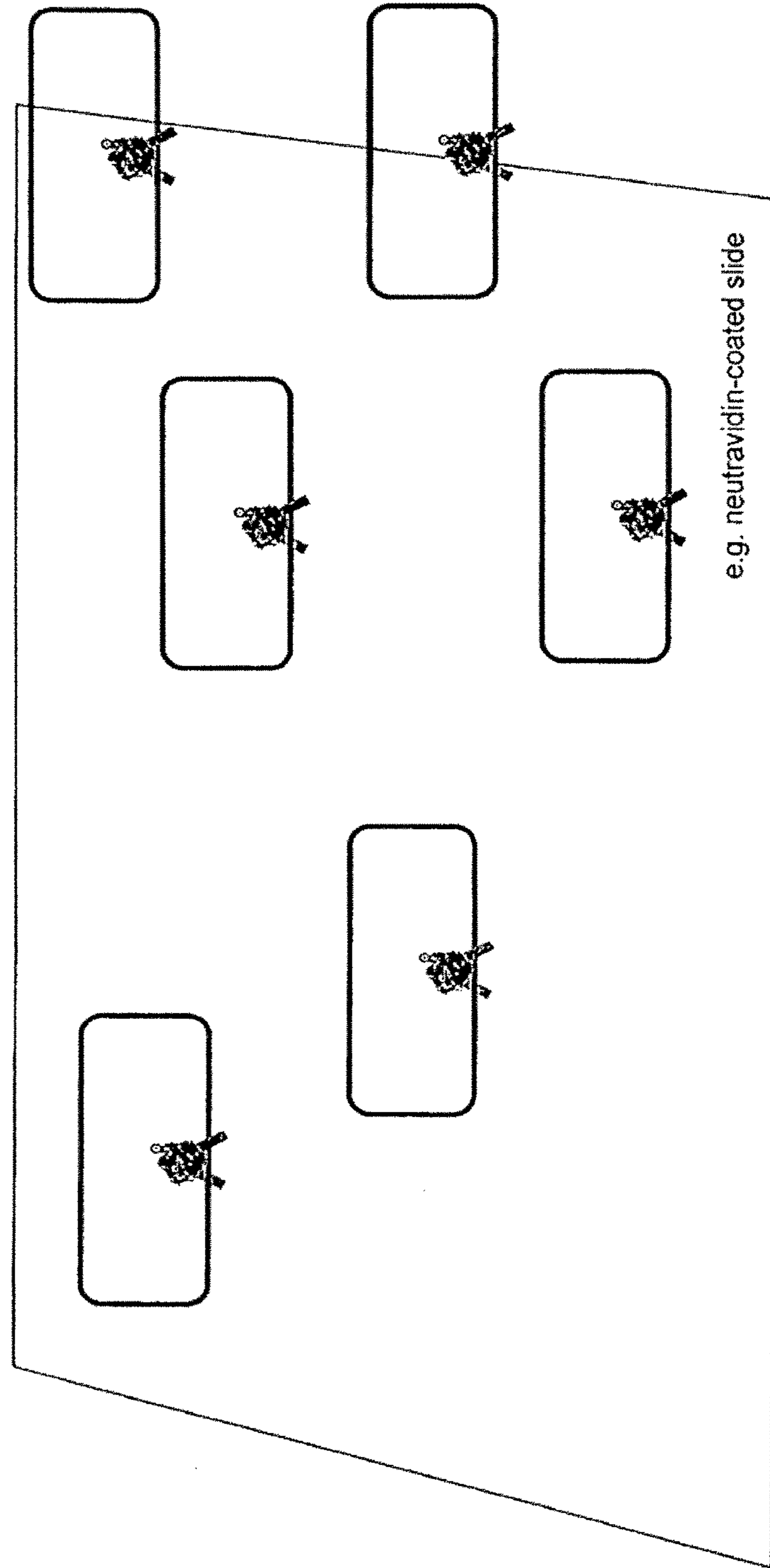


FIG. 22

**POLYMERASE-NUCLEIC ACID COMPLEX****CROSS-REFERENCE TO RELATED APPLICATIONS**

This application is a continuation of U.S. application Ser. No. 11/377,136, filed Mar. 15, 2006, now U.S. Pat. No. 9,045,798 issued Jun. 2, 2015, which is a continuation of U.S. application Ser. No. 11/118,031, filed Apr. 29, 2005, now U.S. Pat. No. 7,462,452 issued Dec. 9, 2008, which claims priority to U.S. Provisional Application No. 60/567,202, filed Apr. 30, 2004, which applications are incorporated herein by reference for all purposes.

**REFERENCE TO A "SEQUENCE LISTING," A TABLE, OR A COMPUTER PROGRAM LISTING APPENDIX SUBMITTED AS AN ASCII TEXT FILE**

The Sequence Listing written in file -34-7.TXT, created on Nov. 25, 2014, 4,096 bytes, machine format IBM-PC, MS-Windows operating system, is hereby incorporated by reference in its entirety for all purposes.

**STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT**

The research embodied within the present application was funded in-part by the Federal Government in research grant number 1 P01 HG003015. The government may have certain rights in this application.

**BACKGROUND OF THE INVENTION**

The primary sequences of nucleic acids are crucial for understanding the function and control of genes and for applying many of the basic techniques of molecular biology. In fact, rapid DNA sequencing has taken on a more central role after the goal to elucidate the entire human genome has been achieved. DNA sequencing is an important tool in genomic analysis as well as other applications, such as genetic identification, forensic analysis, genetic counseling, medical diagnostics, and the like. With respect to the area of medical diagnostic sequencing, disorders, susceptibilities to disorders, and prognoses of disease conditions can be correlated with the presence of particular DNA sequences, or the degree of variation (or mutation) in DNA sequences, at one or more genetic loci. Examples of such phenomena include human leukocyte antigen (HLA) typing, cystic fibrosis, tumor progression and heterogeneity, p53 proto-oncogene mutations, and ras proto-oncogene mutations (see, Gyllensten et al., PCR Methods and Applications, 1:91-98 (1991); U.S. Pat. No. 5,578,443, issued to Santamaria et al.; and U.S. Pat. No. 5,776,677, issued to Tsui et al.).

Various approaches to DNA sequencing exist. The dideoxy chain termination method serves as the basis for all currently available automated DNA sequencing machines, whereby labeled DNA elongation is randomly terminated within particular base groups through the incorporation of chain-terminating inhibitors (generally dideoxynucleoside triphosphates) and size-ordered by either slab gel electrophoresis or capillary electrophoresis (see, Sanger et al., *Proc. Natl. Acad. Sci.*, 74:5463-5467 (1977); Church et al., *Science*, 240:185-188 (1988); Hunkapiller et al., *Science*, 254:59-67 (1991)). Other methods include the chemical degradation method (see, Maxam et al., *Proc. Natl. Acad.*

*Sci.*, 74:560-564 (1977), whole-genome approaches (see, Fleischmann et al., *Science*, 269:496 (1995)), expressed sequence tag sequencing (see, Velculescu et al., *Science*, 270 (1995)), array methods based on sequencing by hybridization (see, Koster et al., *Nature Biotechnology*, 14:1123 (1996)), and single molecule sequencing (SMS) (see, Jett et al., *J. Biomol. Struct. Dyn.* 7:301 (1989); Schecker et al., *Proc. SPIE-Int. Soc. Opt. Eng.* 2386:4 (1995)).

There have been several improvements in the dideoxy chain termination method since it was first reported in the mid-1980's with enhancements in the areas of separating technologies (both in hardware formats & electrophoresis media), fluorescence dye chemistry, polymerase engineering, and applications software. The emphasis on sequencing the human genome with a greatly accelerated timetable along with the introduction of capillary electrophoresis instrumentation that permitted more automation with respect to the fragment separation process allowed the required scale-up to occur without undue pressure to increase laboratory staffing. However, despite such enhancements, the reductions in the cost of delivering finished base sequence have been marginal, at best.

In general, present approaches to improve DNA sequencing technology have either involved: (1) a continued emphasis to enhance throughput while reducing costs via the dideoxy chain termination method; or (2) a paradigm shift away from the dideoxy chain termination method to alternative approaches that do not involve molecular sizing by electrophoretic means.

Although several non-sizing DNA sequencing methods have been demonstrated or proposed, all are limited by short read lengths. For example, matrix-assisted laser desorption/ionization (MALDI) mass spectrometry, which separates DNA fragments by molecular weight, is only capable of determining about 50 nucleotides of DNA sequence due to fragmentation problems associated with ionization. Other non-sizing sequencing methods depend on the cyclic addition of reagents to sequentially identify bases as they are either added or removed from the subject DNA. However, these procedures all suffer from the same problem as the classical Edman degradation method for protein sequencing, namely that synchronization among molecules decays with each cycle because of incomplete reaction at each step. As a result, current non-sizing sequencing methods are unsuitable for sequencing longer portions of DNA.

As such, there is a need for more effective and efficient methods of non-sizing DNA sequencing. The present invention satisfies this and other needs.

**SUMMARY OF THE INVENTION**

The present invention provides compositions, methods and apparatus for single molecule DNA sequencing that can be performed, e.g., in a two-electrode chamber such as a microtiter plate fitted with two electrodes. As such, many different single DNA molecules (e.g., over two-hundred) can be sequenced simultaneously in a single well at a rate of about 10 to about 200 nucleotides per second per molecule and at read lengths of 5-20 kilobases (kb) or more. In addition, the compositions, methods and apparatus of the present invention reduce the cost of sequencing as compared to other long read approaches due to the high degree of multiplexing and the substitution of microtiter plates for expensive micro- or nano-fabricated devices. As such, field-switch sequencing of the present invention provides long-read high-throughput sequencing with sufficient resolution for single molecule detection.

In one embodiment, the present invention provides a particle, comprising:

- a particle surface; at least one nucleotide phosphate (NP) attached to the particle surface via a linker; and
- a dye, wherein the particle carries a charge.

In certain other embodiments, the present invention provides a method for sequencing a nucleic acid, comprising:

- (a) immobilizing a plurality of complexes comprising a target nucleic acid, a primer nucleic acid, and a polymerase onto a surface;
- (b) contacting the surface with a plurality of charged particles comprising a nucleotide phosphate by applying an electric field;
- (c) reversing the electric field to transport unbound charged particles away from the surface; and
- (d) detecting the incorporation of the nucleotide phosphate into a single molecule of the primer nucleic acid.

In yet another embodiment, the present invention provides a sequencing chamber, comprising:

- a first electrode on one side of the chamber;
- a second electrode on another side of the chamber; and
- a spacer separating the first electrode and second electrode.

Other objects, features, and advantages of the present invention will be apparent to one of skill in the art from the following detailed description and figures.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A-F show various particle embodiments of the present invention. FIG. 1A illustrates a NP attached to the particle via a linker, wherein the dye is on the surface of the particle. FIG. 1B illustrates at least two NPs linked to a particle, wherein the dye is on the surface of particle. FIG. 1C illustrates a NP linked to a particle, wherein the dye is embedded in the particle. FIG. 1D illustrates a particle having a NP linked to a particle, and a dye is linked to the particle. FIG. 1E illustrates a trifunctional linker wherein the dye, NP and particle are attached. FIG. 1F shows a NP on one end of the linker and a dye on the other end of a linker.

FIGS. 2A-C show a nucleic acid nanoparticle configuration of the present invention FIG. 2A shows a particle comprising multiple hairpin structures. FIG. 2B shows a single hairpin unit (U-shaped line; basepairs schematically indicated by horizontal line pattern) with 4 intercalated, covalently attached SYBR dyes (black rectangles). FIG. 2C shows an example connector with carboxylate functionality for conjugation to amino-modified bases on the DNA.

FIG. 3 shows a DNA particle (a 2.7-kb pUC plasmid) stained with ~300 molecules each of YOYO-1 or TOTO-1. Samples (1 million particles per uL) were adsorbed to poly-L-lysine coated glass coverslips and the DNA was imaged by objective-type TIR. Excitation power was 1 mW spread over a 100 um diameter spot. Laser wavelengths were: YOYO-1 at 488 nm, TOTO-1 at 514 nm and TOTO-3 at 647 nm. Exposures were 2 msec using a Roper MicroMax 512 CCD camera, and pixel binning was set to 9x9. Signals are indicated for selected particles. Signal to noise (SN) was calculated as the ratio (S-B)/stdev, where S=the integrated signal of a single particle, B=average background, stdev=standard deviation of background pixels.

FIGS. 4A-B show a cycle in the field-switch sequencing method of the present invention. FIG. 4A illustrates the accumulation of negatively-charged particles (white and black circles) above immobilized polymerase-DNA complexes on a positively-charged indium-tin oxide (ITO) electrode (bottom rectangle). Particles bound by polymerases

are shown (black circles). FIG. 4B illustrates the movement of unbound particles away from the ITO electrode when the electric field is reversed. The ITO electrode surface is illuminated by total internal reflection (arrows) and the particles retained by the polymerase-DNA complexes are imaged without interference from unbound particles which have moved away from the surface

FIGS. 5A-B. FIG. 5A shows a diagram of a circular template that is permanently associated with the anchored polymerase, while still being able to slide through the DNA binding groove to permit primer extension. The tunnel formed by polymerase immobilization is roughly the same dimension as a DNA sliding clamp. FIG. 5B shows the structure of a polymerase. The template strand feeds into the DNA binding cleft between the biotin loops (SEQ ID NO:1), trapping it between the enzyme and the immobilization surface. The primer strand, which is extended at its 3'-end by polymerization of dNTPs, exits the polymerase with the template strand as shown.

FIGS. 6A-B. FIG. 6A shows an example of random x-y coordinates for immobilized polymerases, as determined by a computer model. In the simulation, immobilized polymerases closer than 2 microns were treated as being optically unresolvable. Proteins at least 2 microns apart are marked as white squares (they can be sequenced), whereas proteins closer than 2 microns are marked as small circles (too close to sequence). The simulated results show 204 resolvable immobilized proteins among 249 total immobilized proteins. FIG. 6B shows the total number (thin line) and resolvable number (thick line) of immobilized proteins from one simulation run. The results indicate that a maximum of about 350 resolvable proteins is achieved after about 800 total proteins have been immobilized. As the total number of protein increases above 800, the number of resolvable proteins decreases due to overcrowding.

FIG. 7 shows a two-electrode sequencing chamber of the present invention. The bottom and top electrodes are separated by an epoxy (SU-8) spacer film 10 microns thick, as indicated.

FIGS. 8A-D. FIG. 8A shows a "dumbbell" template for production of hairpin particles by rolling circle replication. The 20 nucleotide loop (bottom) is complementary to the primer (circle with arrow). Adenosine bases in the template (squares) allow incorporation of thymidine or uracil bases in the product strand; substituting amino-modified dUTP for dTTP provides amine functionality for attachment of SYBR-dNTP moieties. FIG. 8B shows example oligonucleotides for template construction; oligos 1 and 2 (SEQ ID NOS:2 and 3) are ligated into the dumbbell, and primer 1 (SEQ ID NO:4) is used for rolling circle replication by polymerases capable of strand-displacement polymerization. FIG. 8C shows the rolling circle product, comprising tandem hairpins. FIG. 8D shows secondary structures predicted by mFold software (for a single hairpin unit (SEQ ID NO:5) and for a 6-hairpin particle).

FIGS. 9A-B. FIG. 9A shows ligation of oligos 1 and 2; the resulting "dumbbell" template is purified for use in rolling circle replication by excising the indicated band from the gel. FIG. 9B shows the rolling circle replicated products made with dTTP (lane 3) or aha-dUTP (lane 4; Molecular Probes). Lanes 1-2 are size markers. Lanes 5-9 are enzyme digested particles, showing reduction to 6-hairpin units.

FIG. 10 shows deprotection of BOC-PEG8-amine (Poly-pure), performed before coupling to a nucleotide.

FIG. 11 shows coupling of the deprotected PEG8 diamine (from FIG. 10) to the terminal phosphate group of dATP.

## 5

FIG. 12 shows coupling of SYBR-101 NHS (Molecular Probes) to cysteine.

FIG. 13 shows activation of dCTP-PEG8-amine with SIA. This converts the amine to iodoacetamide functionality for subsequent reaction with a thiol.

FIG. 14 shows the synthesis of SYBR-dCTP with carboxylate functionality contributed by the connector moiety (cysteine).

FIG. 15 shows the electrokinetic properties of 5 simulated particle types: a labeled dNTP not attached to a nanoparticle (curve 1); a DNA dendrimer (curve 2); a 4 kb circular DNA (curve 3); a 2 kb circular DNA (curve 4); and a 55 nm silica-DNA particle having about 50 each of 120-base Y-branched oligonucleotides attached to ~20 nm diameter silica particles (curve 5).

FIGS. 16A-D illustrates various features of a polymerase-nucleic acid complex of the present invention.

FIG. 17 illustrates an anchor embodiment of the present invention.

FIG. 18 illustrates a nucleic acid sample preparation of the present invention.

FIG. 19 illustrates a nucleic acid sample preparation of the present invention.

FIG. 20 illustrates a nucleic acid sample preparation of the present invention.

FIG. 21 illustrates a single molecule isolation embodiment of the present invention.

FIG. 22 illustrates a single molecule bound to a cover slip.

## DETAILED DESCRIPTION OF THE INVENTION

### I. Definitions

As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

The term “oligonucleotide” as used herein includes linear oligomers of nucleotides or analogs thereof, including deoxyribonucleosides, ribonucleosides, and the like. Typically, oligonucleotides range in size from a few monomeric units, e.g., 3-4, to several hundreds of monomeric units. Whenever an oligonucleotide is represented by a sequence of letters, it will be understood that the nucleotides are in 5'-3' order from left to right and that “A” denotes deoxyadenosine, “C” denotes deoxycytidine, “G” denotes deoxyguanosine, and “T” denotes thymidine, unless otherwise noted.

The term “nucleotide” as used herein refers to a phosphate ester of a nucleoside, e.g., mono-, di-, tri-, tetra-, penta-, hexaphosphate esters, and the like, wherein the most common site of esterification is the hydroxyl group attached to the C-5 position of the pentose. Although in many instances throughout this specification “NTP” is referred to, any of the phosphate esters (e.g., mono-, di-, tri-, tetra-, penta-, hexaphosphate esters) can be used interchangeably. Nucleosides also include, but are not limited to, synthetic nucleosides having modified base moieties and/or modified sugar moieties, e.g., described generally by Scheit, *Nucleotide Analogs*, John Wiley, N.Y. (1980). Suitable NTPs include both naturally occurring and synthetic nucleotide triphosphates, and are not limited to, ATP, dATP, CTP, dCTP, GTP, dGTP, TTP, dTTP, UTP, and dUTP. Preferably, the nucleotide triphosphates used in the methods of the present invention are selected from the group consisting of dATP, dCTP, dGTP, dTTP, dUTP, and combinations thereof.

The term “primer nucleic acid” refers to a linear oligonucleotide, which specifically anneals to a unique target

## 6

nucleic acid sequence and allows for synthesis of the complement of the target nucleic acid sequence.

The phrase “target nucleic acid” refers to a nucleic acid or polynucleotide whose sequence identity or ordering or location of nucleosides is to be determined using the methods described herein.

The phrase “sequencing a nucleic acid,” in reference to a target nucleic acid, includes determination of partial as well as full sequence information of the target nucleic acid. That is, the term includes sequence comparisons, fingerprinting, and like levels of information about a target nucleic acid, as well as the express identification and ordering of nucleosides, usually each nucleoside, in a target nucleic acid. The term also includes the determination of the identification, ordering, and locations of one, two, or three of the four types of nucleotides within a target nucleic acid.

The term “linker” refers to a phosphate linker moiety attached to the terminal phosphate (e.g., gamma-phosphate group of an NTP) of a nucleotide phosphate. The linker can contain carbon, oxygen, phosphorous, sulfur or hydrogen atoms. The linker is used to attach the nucleotide phosphate (e.g., NTP) to a particle directly or via a dye.

The term “unit” in reference to a nanoparticle, describes a double-stranded or single-stranded nucleic acid or analog thereof. Individual units are joined together into larger particles by attachment to a common core particle (e.g. a silica nanoparticle), by attachment to each other by complementary base pairing or covalent linking, or by being arrayed along a single strand as tandem direct repeat or inverted repeat sequences.

The term “haripin unit” is a single-strand inverted repeat sequence capable of self-hybridizing into a single stem-loop structure or into multiple stem-loops. Base pairing along the double-stranded stems may be either perfectly complementary or may contain mismatches.

### II. Overview

In certain aspects, the present invention provides methods, compositions and apparatus for field-switch sequencing, wherein nucleotide phosphates (e.g., NTPs) such as deoxyribonucleotide triphosphates (e.g., dNTPs) are attached to nanoparticles. In one embodiment, the terminal phosphate (e.g.,  $\gamma$ -phosphate group of a NTP) is tethered via a linker to the surface of the nanoparticle. In a preferred embodiment, the linker comprises polyethylene glycol (PEG) attached to for example, a trifunctional group for connecting together the dye, the nucleotide and the particle. In another preferred embodiment, the NTP is first linked to a dye, and the NTP-dye compound is then attached to a particle. In certain instances, several hundred NTP-dyes are coupled to the surface of a nanoparticle; for example, given 100% coupling efficiency between a NTP-dye and an amino-DNA, a particle comprising 18 haripin units, wherein each unit contains 17 amino-modified bases, has  $17 \times 18 = 306$  NTP-dyes. Exceptionally bright fluorescence from these nanoparticles enables a charged-couple device (CCD) camera to image about 200-300 individual optically-resolved particles simultaneously with millisecond exposure times. In addition to improved detectability, the nanoparticles are also capable of carrying a substantial electric charge. In certain preferred instances, fluorescence and electric charge are elements of the field-switch sequencing particles of the present invention.

### III. Nanoparticles

Advantageously, the particles of the present invention are used to carry or transport the substrate nucleotide phos-

phates (NPs) to a polymerase (e.g., immobilized polymerase). Each particle has at least one NP, preferably at least two NPs and more preferably a plurality of NPs associated therewith. Preferably, the NP is attached to the particle via a linker. In certain instances, the present invention provides a particle, comprising: a particle surface; at least one nucleotide phosphate attached to the particle surface via a linker; and a dye, wherein the particle carries a charge.

The nucleotide phosphate is for example, a nucleotide diphosphate, a nucleotide triphosphate (NTP), a nucleotide tetraphosphate, a nucleotide pentaphosphate or a nucleotide hexaphosphate. Preferably, a nucleotide phosphate is a nucleotide triphosphate such as a NTP or dNTP.

In certain instances, a label (e.g., a fluorescent dye, a mass tag, a chromogenic labels, and the like) is associated with the particle. Preferably, the label or dye is associate or attached to the surface of the particle. In other aspects, the dye is embedded in the particle. In still other preferred aspects, the dye is attached to one end of the linker, wherein the NP is attached to the other end of the linker.

FIG. 1 illustrates various embodiments of the particles of the present invention. FIG. 1A illustrates a NP attached to the particle via a linker, and wherein the dye is on the surface of the particle. FIG. 1B illustrates at least two NPs linked to a particle, wherein the dye is on the surface of particle. FIG. 1C illustrates a NP linked to a particle, wherein the dye is embedded in the particle. FIG. 1D illustrates a particle having a NP linked to a particle, and a dye is linked to the particle. FIG. 1E illustrates a trifunctional linker wherein the dye, NP and particle are attached via a linker. FIG. 1F shows a NP on one end of the linker and a dye on the other end of a linker. The dye intercalates into the nucleic acid nanoparticle. In certain embodiments, a linker connects the terminal phosphate of the NP to the particle, and a second linker connects the dye to the particle. Preferred linkers comprise backbone chains of carbon, nitrogen, oxygen, sulfur, phosphorus and combinations thereof. In another embodiment, a the NP linker and the dye linker are connected to a common linker, and the common linker is connected to the particle. In a preferred embodiment, the common linker is a trifunctional moiety, for example an amino acid having an amino group, a carboxylate group and a side-chain functional group. In a most preferred embodiment, the common linker is cysteine, aspartate, glutamate or lysine.

A wide variety of particles are suitable for the present invention. These particles include, but are not limited to, a nucleic acid nanoparticle, a nucleic acid analog particle, a semiconductor nanoparticle, a nanocrystal, a silicon nanocrystal, a metallic nanoparticle, a silica-shell nanoparticle, a dye-doped silica-shell nanoparticle, a nanoparticle made of an organic polymer, a dendrimer, a light-scattering nanoparticle, an electroactive nanoparticle, a coated nanoparticle, a lipid nanoparticle, a protein nanoparticle, a fluorescent protein, and combinations thereof. In certain preferred aspects, the particle is a nucleic acid nanoparticle.

Various nucleic acid nanoparticles are useful in the present invention. These nucleic acid nanoparticles include, but are not limited to, a hairpin structure consisting of tandem inverted repeat DNA sequences, a DNA nanoparticle, a RNA nanoparticle, a silica-DNA nanoparticle, a nucleic acid analog nanoparticle, a DNA nanoparticle comprising multiple complementary and non-complementary strands of single stranded DNA, an RNA nanoparticle comprising multiple complementary and non-complementary strands of single stranded RNA, and a combination thereof. Preferably, the nucleic acid particle is a hairpin structure consisting of tandem inverted repeat DNA sequences.

FIGS. 2A-C show a nucleic acid nanoparticle configuration of the present invention. As shown therein, a SYBR-dNTP is intercalated into and covalently attached to, double-stranded DNA comprising a hairpin-type particle. FIG. 2A shows a particle comprising multiple hairpin structures formed from tandem inverted repeat sequences along a single strand. FIG. 2B shows a single hairpin unit (U-shaped line; basepairs schematically indicated by horizontal line pattern) with 4 intercalated, covalently attached SYBR dyes (black rectangles). A tri-functional connector such as an amino acid residue, joins the dye both to the DNA and to a phosphate-linked dNTP. FIG. 2C is an example of a trifunctional connector with carboxylate functionality for conjugation to an amino-modified bases on the DNA.

In certain embodiments, other particle types are suitable for use in the present invention. These particles include, but are not limited to, 20 nm polystyrene spheres (Molecular Probes); semiconductor nanocrystals (Quantum Dot Corporation); RLS metal particles (Invitrogen); dye-doped silica shell nanoparticles; silicon crystal nanoparticles; and DNA dendrimers (Genisphere)). In certain other aspects, nanoparticles made of DNA are useful (Tapeç et al., *J. Nanosci. Nanotechnol.*, 2:405 (2002); Belomoin et al., *App. Phys. Lett.*, 80:841 (2002)). In certain aspects, DNA intercalating dyes from Molecular Probes (e.g., YOYO, JOJO, BOBO, POPO, TOTO, LOLO, SYBR, SYTO, SYTOX, PicoGreen, OliGreen etc.) are useful to intercalate the linker having the NP attached thereto with the particle. As a result, particles containing a plurality dyes (e.g., several hundred) are bright enough for millisecond detection by a CCD camera. DNA provides a structured environment that allows intercalating dyes to pack densely (e.g., 1 dye:10 base pairs) into nanoscale volumes without being quenched. In fact, fluorescence is enhanced several hundred-fold upon intercalation. In addition to high dye density, DNA particles naturally provide: 1) high charge density; 2) facile self-assembly of defined structures; and 3) a flexible particle surface for improved thermal motion of surface-attached dNTPs. Moreover, dye intercalation stabilizes the double helix, as evidenced by increasing the  $T_m$  from 65° C. to 95° C. upon intercalation of 6 YOYO dyes into a 30-mer dsDNA oligonucleotide (Bjorndal et al., *Biopolymers*, 65:40 (2002)).

In certain embodiments, the particle surface comprises a functional group. The functional group can be a reactive functional group. Suitable functional groups include, but are not limited to, a haloacetyl group, an amine, a thiol, a phosphate, a carboxylate, a hydrazine, a hydrazide an aldehyde or a combination thereof. Other functional groups include groups such as a reactive functionality (A) or a complementary group (B) set forth in Table I below.

Preferably, the at least one nucleotide phosphate such as a dNTP is attached to the particle surface via a linker. The linker is a combination of atoms having stable chemical bonds, and optionally includes single, double, triple, aromatic carbon-carbon bonds, as well as carbon-nitrogen bonds, nitrogen-nitrogen bonds, carbon-oxygen bonds and carbon-sulfur bonds. The linker typically includes ether, thioether, carboxamide, sulfonamide, urea, urethane or hydrazine moieties. Preferably, the linker has 1-50 nonhydrogen atoms selected from the group of C, N, O and S; and are composed of any combination of ether, thioether, amine, ester, carboxamide, sulfonamide, hydrazide bonds and aromatic or heteroaromatic bonds. Preferably, the linker is a combination of single carbon-carbon bonds and carboxamide or ether bonds. Examples of the linker include substituted or unsubstituted polymethylene, arylene, alkylarylene or arylenealkyl.

In a preferred embodiment, the linker is an aliphatic chain containing a backbone of 4-50 carbon atoms. The carbon backbone may be interspersed at one or more intervals with a non-carbon heteroatom. The heteroatoms, which may be the same or different are N, O, or S. Where the heteroatom is nitrogen, it is optionally substituted with one or more alkyl substituents having 1-6 carbon atoms, which may be the same or different. In certain aspects, the linker incorporates quaternary nitrogens, which confer cationic charges on the particles of the invention.

In certain preferred aspects, the linker is a polyethylene glycol (PEG) linker. The PEG linker is a  $-(OCH_2CH_2)_n-$  polymer, wherein n is about 1-50, preferably about 4-30 and more preferably about 6-20. The linker attaches to the nucleotide phosphate on one end, and a dye on the other end. The attachment of the linker to the nucleotide phosphate is via complementary functional groups. Similarly, the attachment of the linker to a dye is via complementary functional groups. In another aspect, the linker can be attached directly to the particle without an intervening dye. In a preferred embodiment, both the dye and the NTP are covalently attached to the particle. For example, the dye and NTP are covalently attached via a linker (bi- or trifunctional) to form a dye-NTP moiety, and the dye-NTP moiety is covalently attached to the particle.

Selected examples of reactive functionalities useful for the attaching the linker to the nucleotide phosphate and the linker to a dye or the linker directly to the particle surface are shown in Table I, wherein the bond "C" results from such a reaction. Those of skill in the art will know of other bonds suitable for use in the present invention.

TABLE I

Reactive functionality (A)	Complementary group (B)	The resulting bond (C)
activated esters	amines/anilines	carboxamides
acrylamides	thiols	thioethers
acyl azides	amines/anilines	carboxamides
acyl halides	amines/anilines	carboxamides
acyl halides	alcohols/phenols	esters
acyl nitriles	alcohols/phenols	esters
acyl nitriles	amines/anilines	carboxamides
aldehydes	amines/anilines	imines
aldehydes or ketones	hydrazines	hydrazones
aldehydes or ketones	hydroxylamines	oximes
alkyl halides	amines/anilines	alkyl amines
alkyl halides	carboxylic acids	esters
alkyl halides	thiols	thioethers
alkyl halides	alcohols/phenols	ethers
alkyl sulfonates	thiols	thioethers
alkyl sulfonates	carboxylic acids	esters
alkyl sulfonates	alcohols/phenols	ethers
anhydrides	alcohols/phenols	esters
anhydrides	amines/anilines	carboxamides/imides
aryl halides	thiols	thiophenols
aryl halides	amines	aryl amines
aziridines	thiols	thioethers
boronates	glycols	boronate esters
carboxylic acids	amines/anilines	carboxamides
carboxylic acids	alcohols	esters
carboxylic acids	hydrazines	hydrazides
carbodiimides	carboxylic acids	N-acylureas or anhydrides
diazoalkanes	carboxylic acids	esters
epoxides	thiols (amines)	thioethers (alkyl amines)
epoxides	carboxylic acids	esters
haloacetamides	thiols	thioethers
haloplatinate	amino	platinum complex
haloplatinate	heterocycle	platinum complex
halotriazines	amines/anilines	aminotriazines
halotriazines	alcohols/phenols	triazinyl ethers
imido esters	amines/anilines	amidines
isocyanates	amines/anilines	ureas

TABLE I-continued

	Reactive functionality (A)	Complementary group (B)	The resulting bond (C)
5	isocyanates	alcohols/phenols	urethanes
	isothiocyanates	amines/anilines	thioureas
	maleimides	thiols	thioethers
	phosphoramidites	alcohols	phosphite esters
	silyl halides	alcohols	silyl ethers
	sulfonate esters	amines/anilines	alkyl amines
10	sulfonyl halides	amines/anilines	sulfonamides

In still other embodiments, the particle can further comprise an oligonucleotide or a plurality of oligonucleotides. Due to the highly negatively charged groups, in certain instances, the oligonucleotide imparts a negative charge onto the particle. In other instance, the oligonucleotide can be used as an appendage to attach other moieties for example, attach a dye or an NP (e.g., NTP), to the particle. The NTP can be attached to the oligonucleotides via covalent bonding, non-covalent bonding, via ionic bonding, hydrophobic interactions or a combination thereof.

In another aspect, the particle can be used as a scaffold to attach analytes for example, small molecules (e.g., LIGANDS), for specific interactions with nucleic acids, carbohydrates, lipids, proteins, antibodies, or other ligand binders (TARGETS) bound to the chamber surface. In yet another embodiment, ligands can be attached to the chamber surface and targets can be attached to the particles. Ligands and targets can be attached to particles or the chamber surface by covalent bonding, non-covalent bonding, via ionic bonding, hydrophobic interactions or a combination thereof.

A wide variety of dyes are suitable for use in the present invention. The dye can be associated or attached to the surface of the particle. In other aspects, the dye is embedded into the particle. In still other aspects, the dye is attached to one end of the linker, wherein the NP is associated or attached to the other end of the linker. The dye is preferably a fluorescent dye, or a plurality of fluorescent dyes. Suitable dyes include, but are not limited to, YOYO-1, JOJO-1, LOLO-1, YOYO-3, TOTO, BOBO-3, SYBR, SYTO, SYTOX, PicoGreen, OliGreen, and combinations thereof. Other dyes include, thiazole orange, oxazole yellow, or non-intercalating dyes such as fluorescein, rhodamine, cyanine or coumarin based dyes, and combinations thereof. The dye can be an intercalating dye or a nonintercalating dye. In certain instances, the dye molecules are attached to a plurality of oligonucleotides via covalent bonding. In certain other instances, the dye molecules are attached to the plurality of oligonucleotides via ionic bonding.

Other suitable dyes include, but are not limited to, 4-acetamido-4'-isothiocyanatostilbene-2,2'-disulfonic acid; acridine and derivatives: acridine, acridine isothiocyanate; 5-(2'-aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS); 4-amino-N-[3-vinylsulfonyl]phenyl]naphthalimide-3,5 disulfonate; N-(4-anilino-1-naphthyl)maleimide; anthranilamide; BODIPY; Brilliant Yellow; coumarin and derivatives: coumarin, 7-amino-4-methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcoumarin (Coumaran 151); cyanine dyes; cyanosine; 4',6-diaminidino-2-phenylindole (DAPI); 5',5''-dibromopyrogallol-sulfonaphthalein (Bromopyrogallol Red); 7-diethylamino-3-(4'-isothiocyanatophenyl)-4-methylcoumarin; diethylenetriamine pentaacetate; 4,4'-diisothiocyanatodihydro-stilbene-2,2'-disulfonic acid; 4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid; 5-[dimethylamino]naphthalene-

1-sulfonyl chloride (DNS, dansylchloride); 4-dimethylaminophenylazophenyl-4'-isothiocyanate (DABITC); eosin and derivatives: eosin, eosin isothiocyanate, erythrosin and derivatives: erythrosin B, erythrosin, isothiocyanate; ethidium; fluorescein and derivatives: 5-carboxyfluorescein (FAM), 5-(4,6-dichlorotriazin-2-yl)amino-fluorescein (DTAF), 2',7'-dimethoxy-4'5'-dichloro-6-carboxyfluorescein (JOE), fluorescein, fluorescein isothiocyanate, QFITC, (XRITC); fluorescamine; IR144; IR1446; Malachite Green isothiocyanate; 4-methylumbelliferoneortho cresolphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-phthalaldehyde; pyrene and derivatives: pyrene, pyrene butyrate, succinimidyl 1-pyrene; butyrate quantum dots; Reactive Red 4 (Cibacron™ Brilliant Red 3B-A) rhodamine and derivatives: 6-carboxy-X-rhodamine (ROX), 6-carboxyrhodamine (R6G), lissamine rhodamine B sulfonyl chloride rhodamine (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, sulforhodamine B, sulforhodamine 101, sulfonyl chloride derivative of sulforhodamine 101 (Texas Red); N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); tetramethyl rhodamine; tetramethyl rhodamine isothiocyanate (TRITC); riboflavin; rosolic acid; terbium chelate derivatives; Cy 3; Cy 5; Cy 5.5; Cy 7; IRD 700; IRD 800; La Jolla Blue; phthalocyanine; and naphthalo cyanine.

A wide variety of particles sizes are suitable for the present invention. In certain aspects, the particle has a diameter of about 10 nanometers to about 10 microns. Preferably the particle diameter is about 20 to 700 nanometers, and more preferably, the diameter of about 20 nanometers to about 100 nanometers (e.g., 55 nm).

Detection of the dye labeled particle of the present invention can be performed using a variety of systems and/or devices. Suitable systems and devices include, but are not limited to, an optical reader, a high-efficiency photon detection system, a photodiode, a camera, a charge couple device, an intensified charge couple device, an on-chip multiplication gain charge coupled device, a near-field scanning microscope, a far-field confocal microscope, a microscope that detects wide-field epi-illumination, and a total internal reflection fluorescence microscope.

In certain preferred aspects, the detection device is a total internal reflection fluorescent microscope, using for example, either evanescent wave excitation or surface plasmon excitation. A catalytic event can be detected by one or more cleavage molecules, or while the nucleotide phosphate is resident with the polymerase. In a preferred embodiment, the methods described herein detect the "residence time," or a "resident event" of a nucleotide phosphate on a polymerase, such as within the active site. In certain instances, polymerase catalyzed nucleotide incorporation is synchronized with the application of an electric field. Preferably, detection is carried out by a mechanism that provides low background and enables the detection of single molecules. Various mechanism can be used for example, fluorescence resonance energy transfer, an electron transfer mechanism, an excited-state lifetime mechanism and a ground-state complex quenching mechanism.

In one embodiment, the detection device is a four color single molecule imaging microscope with total internal reflectance (e.g., evanescent wave) excitation and a CCD camera to image a field of view (e.g., a 100  $\mu\text{m}$   $\times$  100  $\mu\text{m}$  field of view). The camera is, e.g., a Micromax 512 $\times$ 512 or a Pentamax intensified CCD from Roper Scientific. The mean background noise with the laser on is 10 electrons. With 2 $\times$ 2 pixel binning and a 2 MHz readout, the total frame readout time is 33 msec, which is compatible with a 100 msec

field-switch sequencing cycle. Single immobilized TAMRA molecules are typically imaged in 80 msec with an SN of ~11.

In certain preferred aspects, the staining of DNA with an intercalating dye give sufficient SN for a nano-scale particle. For example, as shown in FIG. 3, 2.7 kb DNA fragments containing about 300 YOYO-1 or TOTO-1 dye molecules were imaged in 2 msec with SN 186 and 75, respectively. This SN is more than sufficient for unambiguous detection of the particles.

In a further embodiment, a frame transfer CCD is used to capture images (Roper Micromax, I-Pentamax, Cascade). These cameras are capable of a short 1 msec exposure followed by a 40 msec readout. This timing is compatible with the field-switch cycle. Since these cameras are capable of acquiring two images in rapid succession (spaced by 1 msec) followed by an 80 msec readout, two images per cycle are acquired. In one aspect, four color detection is accomplished by a dichroic image splitter (Roper Scientific) mounted in front of a camera.

In certain other instances, one or more dyes or used, in other instances, two or more dyes or used, in other instances, three or more dyes or used and in still other instances, four or more dyes or used. In one embodiment, two dyes are mixed in 4 different ratios within the particles in order to provide 4 different fluorescence signatures (i.e. "colors") using only two dyes. In general, it is not necessary to associate a different dye with each NTP type; rather, two or more dyes can be used in unique dye mixtures to produce the required number of fluorescence signatures. While it is preferred for sequencing to associate a unique fluorescence signature with each nucleotide type attached to the particle, in some applications, 2 or more nucleotide types are attached to the same particle, or different particles having the same fluorescence signature. In yet other embodiments, additional binding moieties are attached to the particles in order to enhance binding between the polymerase and the particle.

Suitable dyes for use in the nanoparticles of the present invention include, without limitation, a family of homodimeric cyanine DNA intercalating dyes from Molecular Probes that cover the visible spectrum, such as YOYO-1 (488/509), JOJO-1 (532/545), LOLO-1 (565/579), and YOYO-3 (612/631), SYBR-101 (488/505) and SYTO-62 (652/676). Given sufficient detection SN, dyes are mixed in various ratios in a single particle such that, for example, different fluorescence spectra are obtained from mixtures of just 2 dyes. In monitoring the fluorescence of individual stained DNA molecules (JOJO-1) over time, an exponential photobleaching decay constant of  $8 \times 10^{-4}$  msec<sup>-1</sup> was determined, meaning that fluorescence decreases 8 parts per 10,000 for each msec of exposure to the excitation laser beam (2.7 W/cm<sup>2</sup>). Particles are preferably illuminated while trapped by a polymerase, and detected.

#### IV. Field Switch Sequencing Methods

In certain embodiments, the present invention provides a method for sequencing a nucleic acid, comprising:

- (a) immobilizing a plurality of complexes comprising a target nucleic acid, a primer nucleic acid, and a polymerase onto a surface;
- (b) contacting the surface with a plurality of charged particles comprising a nucleotide phosphate by applying an electric field;
- (c) reversing the electric field to transport unbound charged particles away from the surface; and

(d) detecting the incorporation of the nucleotide phosphate into a single molecule of the primer nucleic acid.

In certain aspects, the field-switch sequencing methods of the present invention comprise cycled transport of nanoparticle nucleotides between a bottom electrode and a top electrode (FIGS. 4A-B). In one embodiment, the bottom electrode is the glass bottom of a microtiter well coated with electrically-conductive, optically-transparent indium-tin oxide (ITO). In one aspect, about 2 to 1000, preferably 50 to 700, more preferably 200-300 polymerase-DNA complexes are immobilized in the field of view at random positions on the bottom of a well, such that the majority of complexes are optically resolvable from their nearest neighbors. This allows many different molecules (e.g., about 200-300) to be sequenced simultaneously by imaging with a CCD camera (e.g., a 100  $\mu\text{m}$  field).

In certain preferred aspects, the sequencing cycle comprises a wave of charged particles, which is cycled between electrodes by an alternating electric field (E-field). First, charged particles (e.g., anions) are concentrated at the bottom electrode to blanket the immobilized polymerase-DNA complexes. This allows polymerases to bind the correct nucleotides for incorporation into DNA. Next, the E-field is reversed to transport unbound particles away from the surface, leaving only particles retained by the polymerases. With unbound particles now cleared from the surface (e.g., a 200-800 nm distance is sufficient), retained particles are imaged by for example, evanescent wave or surface plasmon excitation with millisecond time resolution while the catalytic reaction is in progress.

FIGS. 4A-B show a cycle in the field-switch sequencing method of the present invention. With reference to FIG. 4A, it is shown that negatively-charged particles (white and black circles) accumulate above immobilized polymerase-DNA complexes on a positively-charged indium-tin oxide (ITO) electrode (bottom rectangle). Particles bound by polymerases are shown (black circles). After the E-field is reversed, FIG. 4B illustrates the movement of unbound particles away from the electrode. The electrode surface is illuminated by total internal reflection (arrows) and the particles retained by the polymerase-DNA complexes are imaged without interference from unbound particles which have moved away from the surface. Each particle comprises at least 1 NP, or at least two NPs, or a plurality of NPs.

In certain aspects, images are acquired before the catalytic reaction is completed because, after incorporation of the nucleotide into DNA, the terminal phosphate (e.g., pyrophosphate) and the attached nanoparticle are released from the enzyme. This completes one sequencing cycle. The timing of E-field switching and image acquisition is dictated by the duration of the catalytic cycle, and ranges from about 1 msec to 1 sec (e.g., 1 msec-100 msec) (see, Levene et al., *Science*, 299:682 (2003)). In certain aspects, the detecting comprises detecting one or more cleavage molecules, either before, during or after the E-field is reversed. In other instances, the detecting comprises detecting the nucleotide phosphate (e.g., dNTP) while associated with the polymerase.

In certain other aspects, sequencing is performed with only 1 image per cycle, taken as soon as the free particles are transported away from the surface. In a preferred embodiment, a second image is taken at the end of each cycle to identify any "background" particles still bound to the electrode. In yet another embodiment, multiple images are taken during each cycle to monitor the release time of individual particles bound to the immobilized polymerases. High nucleotide surface density, flexible tethered nucleotides, and

high particle concentrations at the electrode all have positive effects on polymerase binding kinetics.

In one embodiment, when the field-switch cycle operates at about 10 cycles/sec (i.e., 100 msec period), the sequencing speed is about 10 bases per second. At this speed, a 20 kb DNA molecule is sequenced in about 33 min. In addition, net throughput is significantly enhanced by multiplexing. For example, with an average of one polymerase-DNA complex per 50  $\mu\text{m}^2$  area, there are about 200 optically-resolved complexes in the optical field (100 $\times$ 100  $\mu\text{m}$ ) on the imaged electrode surface. In this embodiment, each well is used once for a period of 33 min. to simultaneously sequence all 200 20 kb DNA molecules (i.e., 4 million bases total). Then, the next 4 million bases are sequenced at a new region of the electrode, and so on. Although it takes about 30-40 days to process an electrode with 1536 individual observation areas, one observation area at a time, the plate is processed 4 times faster (i.e., in 7-10 days) by quadruplexing the instrument optics. Under these conditions, a 1536 well microtiter plate (i.e. one observation area per microtiter well) produces the equivalent of 2 human genomes worth of sequence (i.e., 4 million bases/well $\times$ 1536 wells=6.1 billion bases total).

#### A. Topologically Linked Polymerase-DNA Complexes

In a preferred embodiment, the polymerase-DNA complexes are taught and described in U.S. Patent Publication No. 2005/0042633, published Feb. 24, 2005, and incorporated herein by reference. As described therein, a polymerase-nucleic acid complex (PNAC), comprises: a target nucleic acid and a nucleic acid polymerase, wherein the polymerase has an attachment complex comprising at least one anchor, which at least one anchor irreversibly associates the target nucleic acid with the polymerase to increase the processivity index. As used herein, the term "processivity index" means the number of nucleotides incorporated before the polymerase dissociates from the DNA. Processivity refers to the ability of the enzyme to catalyze many different reactions without releasing its substrate. That is, the number of phosphodiester bonds formed is greatly increased as the substrate is associated with polymerase via an anchor.

##### 1. Polymerase-Nucleic Acid Complex

In one embodiment, the present invention provides a polymerase-nucleic acid complex (PNAC), comprising: a target nucleic acid and a nucleic acid polymerase, wherein the polymerase has an attachment complex comprising at least one anchor, which at least one anchor irreversibly associates the target nucleic acid with the polymerase to increase the processivity index. As used herein, the term "processivity index" means the number of nucleotides incorporated before the polymerase dissociates from the DNA. Processivity refers to the ability of the enzyme to catalyze many different reactions without releasing its substrate. That is, the number of phosphodiester bonds formed using the present invention is greatly increased as the substrate is associated with polymerase via an anchor.

In one embodiment, the processivity index is defined as the number of nucleotides sequenced divided by the number of nucleotides in the template. For example, if the template is 10,000 bases long, and the PNAC sequences 9000 bases, the index is 0.90. Using the PNACs and methods of the present invention, the index is preferably between at least 0.5 to about 1. More preferably, the index is about at least 0.80 to about 1, such as at least 0.80, or at least 0.85, or at least 0.90, or at least 0.95, or 1.0.

Using the PNACs of the present invention, because the target is irreversibly associated with the polymerase, the



number of nucleotides added can be from about 20 to about 100,000, such as about 1000 to about 30,000, such as about 5000 to about 20,000.

FIGS. 16A-D are examples of polymerase nucleic acid complexes (PNACs) of the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art will recognize other variations, modifications, and alternatives.

The polymerase-nucleic complex comprises at least one anchor. In certain aspects, the PNAC will further comprise a primer, which complements a region of the target nucleic acid. As shown in FIG. 16A, the polymerase **101** can have at least one anchor **130** such anchor comprising for example, an amino acid, an epitope, a modified amino acid and the like, for attaching a topological tether. The amino acid i.e., anchor can be for example, a cysteine or a histidine. In certain aspects, the polymerase nucleic acid complex, wherein the nucleic acid **120** is preferably within the active site, comprises at least two anchors. Suitable anchors of the present invention include, but are not limited to, an amino acid, a modified amino acid, a peptide, a histidine tag, a histidine patch, an epitope, and the like. In certain instances, the at least one anchor entraps the target nucleic acid such as by folding back on itself. In other instances, the anchors of the present invention are useful for also attaching a topological tether to the polymerase, or for example, attaching the PNAC to a substrate. In other embodiments, the anchor affixes the PNAC to a support, with or without a topological tether. In certain other embodiments, the polymerase-nucleic complex comprises a topological tether bound to at least two anchors.

As shown in FIG. 16B, an anchor **130** can further comprise other functionalities such as a first member **135** of a first binding pair. A second anchor **140** has a first member **145** of a second binding pair. As shown in FIG. 16C, in certain instances, a topological tether is formed when the first members **135**, **145** are joined by a common member **148**. Alternatively, a topological tether can be formed when the first members **135**, **145** are each joined directly to a support (not shown). A topological tether and at least one anchor can attach via complementary binding pairs. Alternatively, the anchors can attach directly to a substrate without the use of a tether (for example, histidine patches as anchors bound directed to a Ni surface). Suitable complementary binding pairs include, but are not limited to, any haptenic or antigenic compound in combination with a corresponding antibody or binding portion or fragment thereof, nonimmunological binding pairs, receptor-receptor agonist or antagonist, IgG-protein A, lectin-carbohydrate, enzyme-enzyme cofactor, enzyme-enzyme-inhibitor, and complementary polynucleotide pairs capable of forming nucleic acid duplexes.

Exemplary complementary binding pairs include, but are not limited to, digoxigenin and anti-digoxigenin, fluorescein and anti-fluorescein, dinitrophenol and anti-dinitrophenol, bromodeoxyuridine and anti-bromodeoxyuridine, mouse immunoglobulin and goat anti-mouse immunoglobulin, biotin-avidin, biotin-streptavidin, thyroxine and cortisol, histidine patch and Ni-NTA and acetylcholine and receptor-acetylcholine. In certain aspects, the anchor comprises at least one amino acid or an epitope for attaching the topological tether.

As discussed, in certain instances, anchors can comprise an amino acids capable of modification for attachment to a binding member, a tether, a support, and combinations thereof. In one embodiment, a topological tether can attach to two anchors, without intervening binding pairs.

In one aspect, the anchor comprises a biotin moiety. For example, biotin-X nitrilotriacetic acid can be used to covalently attach the biotin moiety to a protein having a free amino group. In turn, this biotin anchor can attach to a streptavidin or a neutravidin binding member, or alternatively, directly to a streptavidin or a neutravidin support.

In another aspect, the topological tether comprises an antibody. In certain embodiments, the topological tether is an antibody that can attach via anchors having complementary binding pairs. For example, the two anchors can be histidine tags, and the tether can be an antibody. In certain aspects, the polymerase-nucleic complex comprises a topological tether anchored to a solid support **150** (see, FIG. 16D).

In certain aspects, the polymerase-nucleic acid attachment complex can be attached to the substrate by providing an anchor such as a polyhistidine tag, that binds to metal. Other conventional means for attachment employ binding pairs. Alternatively, covalent crosslinking agents can be employed such as reagents capable of forming disulfide (S—S), glycol (—CH(OH)—CH(OH)—), azo (—N=N—), sulfone (—S(=O)<sub>2</sub>—), ester (—C(=O)—O—), or amide (—C(=O)—N—) bridges. The covalent bond is for example, an amide, a secondary or tertiary amine, a carbamate, an ester, an ether, an oxime, a phosphate ester, a sulfonamide, a thioether, a thiourea, or a urea.

Selected examples of reactive functionalities useful for the attaching an anchor to the polymerase, a tether to the anchor, or the PNAC to the substrate are shown in Table I, wherein the bond results from such a reaction. Those of skill in the art will know of other bonds suitable for use in the present invention.

In certain aspects, the polymerase can be covalently attached to a support (e.g., coverslip, metal surface, and the like), wherein the polymerase is labeled in vivo with a modified amino acid such as for example, a benzaldehyde derivative of phenylalanine. In one example, the benzaldehyde derivative of phenylalanine is p-acetyl-L-phenylalanine, which can be labeled at specific position(s) in the polymerase. This can be accomplished using organisms (e.g., *E. coli*, yeast) engineered to have an augmented 21-amino acid genetic code capable of inserting p-acetyl-L-phenylalanine at specific codons (see, Lei Wang, Zhiwen Zhang, Ansgar Brock, Peter G. Schultz (2003) *Proc Natl Acad Sci USA* 100:56-61). In one aspect, the polymerase gene of the present invention is engineered to have the appropriate codon or codons at the desired anchor positions, and the corresponding polymerase protein is expressed in the 21-amino acid organism. The expressed polymerase is then purified, mixed with the template DNA, and the resulting PNACs are contacted to a support derivatized with a hydrazine, hydrazone, and the like (e.g., SANH from Solulink Inc). Alternatively, a chemical functionality equivalent to p-acetyl-L-phenylalanine can be attached to the protein at specific or unspecific positions by conjugating SFB (Solulink Inc) to lysine amino acids on the protein. The functionalized protein is attached to the support as above.

In one preferred embodiment, the target nucleic acid is circular DNA. In one aspect, the circular DNA is sequenced by strand displacement synthesis. As is shown in FIG. 18, randomly-sheared fragments of genomic DNA are purified from a sample organism. The DNA **300** is then treated with for example, T4 DNA polymerase, to generate blunt ends and a single “A” nucleotide is added to the 3'-ends with for example, Taq DNA polymerase, and dATP. A mixture of two double-stranded oligonucleotide adaptors **301** and **302** (each with a “T” nucleotide on one 3'-end to complement the “A”

nucleotide on the randomly-sheared fragment) is ligated to the DNA fragments **300** with T4 DNA ligase, wherein the first adaptor **301** is 5'-biotinylated on one strand and the second adaptor **302** is not biotinylated. Whereas the adaptors attach with equal probability to the DNA fragment ends, about half of the ligated DNA molecules will have one biotinylated adaptor and one non-biotinylated adaptor, one quarter will have two biotinylated adaptors, and one quarter will have two non-biotinylated adaptors as shown in FIG. **18**. The desired ligated DNA fragment types, having one biotinylated and one non-biotinylated adaptor, are purified after ligation using gel electrophoresis and streptavidin-coated magnetic beads as follows.

After ligation, DNA fragments in the size range of about 17-23 kb are purified by gel electrophoresis. As shown in FIG. **19**, the purified fragments are bound to streptavidin-coated magnetic beads (Dyna). After binding, the beads are washed to remove unbound DNA. Then the bound DNA is denatured at alkaline pH and the unbiotinylated strands **401** are eluted and the DNA still bound to the beads is discarded. As shown in FIG. **20**, the eluted strands are circularized by hybridizing to a primer oligonucleotide complementary to both adaptors and ligating the two ends of the eluted strand.

FIG. **17** shows a structural model of a PNAC comprising a 9 Degrees North DNA polymerase (parent of Terminator polymerase) **202** and a circular primed DNA template **200**. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art will recognize other variations, modifications, and alternatives. The polymerase **202** comprises anchors **203** and **205** inserted at Terminator amino acid positions K53 and K229, respectively. The anchors are identical in amino acid sequence (LLSKKRSLCCXCTVIVYVTD), wherein the anchor comprises amino acid pa-Phe, which is indicated by "X" in the sequence and by white diamonds **204**, **206**. The pa-Phe amino acids **204**, **206** are shown attached to the support **207**. The circular DNA template **200** is hybridized to a primer **201**. The 5'-end of the primer is indicated **201** and the 3'-end of the primer is hidden in the DNA binding cleft of the protein **202**. The structural model is 1QHT.pdb in the protein database at [www.rcsb.org/pdb/](http://www.rcsb.org/pdb/).

As discussed, the Terminator DNA polymerase can be modified by inserting a 20-amino acid anchor at position K53 and a 20-amino acid anchor at position K229 in the Terminator gene. These two positions straddle the DNA binding cleft as shown in FIG. **17**. As shown therein, each 20-amino acid anchor is engineered to contain at least one p-acetyl-L-phenylalanine (pa-Phe) amino acid near the middle of the anchor (FIG. **17**). The engineered protein is then purified. In one embodiment, to make polymerase nucleic acid complexes, the purified Terminator protein is mixed with a primed single stranded circular DNA template and the mixture is contacted with a support derivatized with hydrazine or hydrazone linkers (Solulink Inc). Optionally, the template DNA contains at least one dUTP base positioned 4-5 bases from the 3'-end of the primer in order to stabilize the polymerase-DNA complex as described (see, Mark Fogg, Laurence Pearl, Bernard Connolly (2002) *Nature Structural Biology* 9:922-927). The polymerase-DNA complex attaches to the support by bond formation between the pa-Phe on the protein and the hydrazine or hydrazone linker on the support. Optionally, the kinetics of bond formation can be increased by concentrating polymerase-DNA complexes on the support surface using an energy field (e.g., electric field, pressure field, magnetic field, and the like). Once the PNAC has formed on the

support, the circular DNA is irreversibly associated with the polymerase as shown in FIG. **17**.

## 2. Immobilization of the PNACs

In certain embodiments, the PNAC arrays of the present invention are immobilized on a support. Preferably, the support (e.g., solid support) comprises a bioreactive moiety or bioadhesive layer. The support can be for example, glass, silica, plastic or any other conventionally material that will not create significant noise or background for the detection methods. The bioadhesive layer can be an ionic adsorbent material such as gold, nickel, or copper, protein-adsorbing plastics such as polystyrene (U.S. Pat. No. 5,858,801), or a covalent reactant such as a thiol group.

The PNAC arrays of the present invention can be immobilized on a support in a random fashion (e.g., random X or Y position coordinates), uniform fashion (e.g., regularly spaced X or Y position coordinates) or a combination thereof. As is shown in FIG. **21**, in one aspect, the PNAC are isolated into single molecule configuration. This single molecule isolation enables efficient attachment of the PNACs to the support. In addition, it allows for efficient single molecule sequencing. Advantageously, the present invention provides single PNACs attached so as to be optically resolvable from their nearest neighbor PNACs. Thus, the PNACs can be analyzed individually without interference from overlapping optical signals from neighboring PNACs. In the present invention, many individual optically resolved PNACs can be sequenced simultaneously.

FIG. **22** is an example of a randomly associated array of PNACs immobilized on a neutravidin-coated slide. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art will recognize other variations, modifications, and alternatives. As shown therein, PNACs are attached or immobilized to a neutravidin-coated slide via an anchor having for example, the first member of a binding pair, wherein the anchor comprises a biotin moiety. In operation, multiple sites can be sequenced with ease.

In yet another example, the PNACs can be attached to the bioadhesive pattern by providing a polyhistidine tag on the polymerase that binds to metal bioadhesive patterns. To create a patterned or random array of a bioadhesive layer, an electron-sensitive polymer such as polymethyl methacrylate (PMMA) coated onto the support is etched in any desired pattern using an electron beam followed by development to remove the sensitized polymer. The holes in the polymer are then coated with a metal such as nickel, and the polymer is removed with a solvent, leaving a pattern of metal posts on the substrate. This method of electron beam lithography provides the very high spatial resolution and small feature size required to immobilize just one molecule at each point in the patterned array. An alternate means for creating high-resolution patterned arrays is atomic force microscopy. A third means is X-ray lithography.

Other conventional means for attachment employ homobifunctional and heterobifunctional crosslinking reagents. Homobifunctional reagents carry two identical functional groups, whereas heterobifunctional reagents contain two dissimilar functional groups to link the biologics to the bioadhesive. A vast majority of the heterobifunctional crosslinking agents contain a primary amine-reactive group and a thiol-reactive group. Covalent crosslinking agents are selected from reagents capable of forming disulfide (S—S), glycol (—CH(OH)—CH(OH)—), azo (—N=N—), sulfone (—S(=O)<sub>2</sub>—), ester (—C(=O)—O—), or amide (—C(=O)—N—) bridges.

A bioresist layer may be placed or superimposed upon the bioadhesive layer either before or after attachment of the biologic to the bioadhesive layer. The bioresist layer is any material that does not bind the biologic. Examples include bovine serum albumin, neutravidin, gelatin, lysozyme, octoxynol, polysorbate 20 (polyethenesorbite monolaurate) and polyethylene oxide containing block copolymers and surfactants (U.S. Pat. No. 5,858,801). Deposition of the layers is done by conventional means, including spraying, immersion and evaporative deposition (metals).

In a preferred embodiment, the polymerase is attached to the ITO permeation layer and stably associated with a DNA template to achieve long sequence reads. The polymerase can be attached to the ITO permeation layer via various linkages including, but not limited to, covalent, ionic, hydrogen bonding, Van der Waals' forces, and mechanical bonding. Preferably, the linkage is a strong non-covalent interaction (e.g. avidin-biotin) or is covalent. In order to permanently associate the DNA template and the polymerase to the ITO permeation layer, an approach that functionally mimics the sliding clamp of a replisome, as described in Shamoo et al., *Cell*, 99:155 (1999), can be used. As shown in FIG. 5, the polymerase-DNA complex is attached to the ITO permeation layer through two biotin modifications on the polymerase binding to streptavidin covalently linked to the permeation layer. This topology irreversibly captures the DNA while still allowing it to slide through the polymerase active site. Circular in form, the DNA (~20 kb) is topologically linked to the immobilized polymerase, and therefore does not dissociate.

#### B. Polymerases

In certain instances, the methods of the present invention employ a DNA polymerase such as DNA polymerase I, II, or III. In certain other instances, suitable polymerases include, but are not limited to, a DNA-dependent RNA polymerase and reverse transcriptase such as an HIV reverse transcriptase. Specific examples include, but are not limited to, T7 DNA polymerase,  $\phi$ 29 DNA polymerase, T5 DNA polymerase, *E. coli* DNA polymerase I, T4 DNA polymerase, T7 RNA polymerase, Taq DNA polymerase, Vent DNA polymerase and Therminator polymerase. Those of skill in the art will know of other enzymes or polymerases suitable for use in the present invention.

In certain aspects, the polymerases useful in the present invention are selected from the A family polymerases or the B family polymerases. DNA-dependent DNA polymerases have been grouped into families, including A, B, X, and others on the basis of sequence similarities. Members of family A, which includes bacterial and bacteriophage polymerases, share significant similarity to *E. coli* polymerase I; hence family A is also known as the pol I family. The bacterial polymerases also contain an exonuclease activity, which is coded for in the N-terminal portion. Family A polymerases include for example, Klenow, Taq, and T7 polymerases. Family B polymerases include for example, the Therminator polymerase, phi29, RB-69 and T4 polymerases.

In certain instances, suitable DNA polymerases can be modified for use in the present invention. These polymerases include, but are not limited to, DNA polymerases from organisms such as *Thermus flavus*, *Pyrococcus furiosus*, *Thermotoga neapolitana*, *Thermococcus litoralis*, *Sulfolobus solfataricus*, *Thermatoga maritima*, *E. coli* phage T5, and *E. coli* phage T4. The DNA polymerases may be thermostable or not thermostable.

In other embodiments, the polymerases include T7 DNA polymerase, T5 DNA polymerase, HIV reverse tran-

scriptase, *E. coli* DNA pol I, T4 DNA polymerase, T7 RNA polymerase, Taq DNA polymerase and *E. coli* RNA polymerase. In certain instances, exonuclease-defective versions of these polymerases are preferred. The efficiency with which  $\gamma$ -labeled NTPs are incorporated may vary between polymerases; HIV-1 RT and *E. coli* RNA polymerase reportedly readily incorporate  $\gamma$ -labeled nucleotide. The polymerase can also be a T7 polymerase. T7 polymerase has a known 3D structure and is known to be processive. In order to operate in a strand-displacement mode, the polymerase requires a complex of three proteins: T7 polymerase+thioredoxin+primase (Chowdhury et al. PNAS 97:12469). In other embodiments, the polymerases can also be HIV RT and DNA Polymerase I.

For Therminator polymerase, protein regions on either side of the DNA binding cleft likely to be conformationally rigid were identified based upon previous studies with RB69 polymerase. Loops of ten amino acids containing a 6xHis sequence at five candidate positions were inserted. Loops inserted at positions K53 and K229 had no deleterious effect on polymerase activity when present either individually or combined.

Suitable covalent coupling methods include, without limitation, a maleimide or thiol-activated permeation layer coupled to specific cysteine amino acids on the polymerase surface, a carboxylate permeation layer coupled to specific lysine amino acids on the polymerase surface, a hydrazine permeation layer coupled to the unnatural amino acid p-acetyl-L-phenylalanine on the polymerase surface, and the like. The latter is particularly useful because of its high coupling specificity, and long reactant shelf life. Given a suitable coupling chemistry, complexes are formed by mixing the polymerase with primed circular DNA and driving them electrically to the electrode surface for covalent coupling. To ensure that most anchored proteins are associated with template, DNA is used at concentrations exceeding the binding constant. Polymerases anchored without DNA are neglected because they have no sequencing activity. In certain aspects, when polymerase attachment is complete, the electric field is reversed to elute linear (e.g., broken) DNA templates, such that the only anchored polymerases capable of generating sequence data are those complexed with circular DNA templates. As shown in FIG. 6A, a simple computer model indicates that polymerase-DNA complexes (e.g., 200-300) can be dispersed randomly in a field (e.g. a 100  $\mu$ m) of view at optically resolvable distances. In certain preferred aspects, polymerase-DNA complexes being optically resolvable is an important feature of the present invention. The number of resolvable complexes decreases at higher densities (FIG. 6B) because of overcrowding. Random dispersion on a surface (e.g., ITO) provides an easy way to isolate single molecules for multiplexed, long-read sequence analysis. In some aspects, polymerase-DNA complexes can be attached in random orientation to the electrode by either covalent or non-covalent interactions; polymerases attached in active orientations are functional for sequencing.

#### V. Field-Switch Cycle and Electrode Design

In one embodiment, the present invention provides an integration of electrophoretic particle control with an ITO electrode and total internal reflection (TIR) optics for field-switch sequencing. This type of setup has been used, e.g., to accelerate antibody-antigen binding kinetics and confine fluorescence detection to the electrode surface (Asanov et al., *Anal. Chem.*, 70:1156 (1998); Liron et al., *Biosens. Bioelectron.*, 17:489 (2002)). Electrophoretic control has

been used (to control hybridization kinetics on DNA chips operating both with faradaic current (Sethi et al., *Clin. Chem.*, 50:443 (2004)) and non-faradaic (capacitive) current (Heaton et al., Proc. Natl. Acad. Sci. USA 98:3701 (2001)). Ultrasensitive fluorescence detection has been demonstrated

on ITO electrodes by imaging single dye molecules with TIR optics (Lu et al., *J. Phys. Chem. B.*, 101:2753 (1997)). Table II shows an example of a field-switch cycle, in which particle transport times were determined by simulation with the computer model described above. In a preliminary step, particles are transported (10  $\mu\text{m}$ ) from the top electrode to the ITO electrode (“ITO”) in 2.4 msec (step #1). Due to a small amount of diffusive broadening in transit, particles accumulate at the ITO over a period of  $\sim 0.1$  msec. Time-critical steps immediately follow, with initiation of polymerase binding marking the start of a new cycle:  $t_o=0.0$  msec (step #2). At  $t=0.3$  msec, the electric field is reversed to transport particles away from the ITO (step #3). The surface is imaged as soon as unbound particles have cleared the evanescent zone:  $t=0.6$  msec (step #4). At 2.8 msec, the particles have accumulated at the top electrode (step #5), where they are held (step #6) until needed for the next cycle (repeat from step #1). In this 100 msec cycle, particles spend  $\sim 95$  msec idle at the top electrode to allow all polymerases to complete their respective catalytic cycles. To maintain electrochemical balance, the electric field alternated rapidly (e.g., 1-10 kHz) at reduced voltages.

TABLE II

The Field-switch Cycle of 55 nm Particles at 1 pN Particle Force.				
Step	Start (msec)	Finish (msec)	Particle Position	Image
1	-2.4	0	Transit to ITO	
2	0.0	0.3	Polymerase binding at ITO surface	
3	0.3	0.6	Transit to top	
4	0.6	2.7	Transit to top	Acquire 2.1 msec
5	2.1	2.8	Collect at top surface	
6	2.8	97.5	Idle at top surface	
7	97.5	99.9	Transit to ITO	
8	99.9	100.0	Collect at ITO surface	
Repeat from Step 1				

**Timing.** In order to correctly trigger the time-critical steps in the beginning of the cycle, it is important to know the particle arrival time at the ITO with at least 0.1 msec certainty (i.e., the arrival time window). In one embodiment, bulk particle arrival is detected by a rise in fluorescence using a fast photodiode detector with low-power illumination. In another embodiment, bulk particle arrival is detected by a current drop as particles passivate the electrode surface. In yet another embodiment, bulk particle arrival is detected by a known time interval between switching the field to transport particles from the top electrode to the bottom electrode.

**True Force on a Particle.** For the methods of the present invention, a particle force limit of 1 pN was arbitrarily set. This requires knowing the real electric force on the particle during operation of the field-switch cycle. While the computer model described above estimates force as a function of electric field strength, some required user inputs may not be known accurately. For example, the true particle charge is difficult to quantify because of uncertainty in the number of particle-associated counter ions, and the true electric field

strength is difficult to quantify due to the steep voltage drop within the electrical double-layers at the electrodes. Fortunately, determination of particle transit times allows calculation of particle terminal velocity (i.e., transit distance/transit time), which yields the electric force as follows: at terminal velocity (attained in nanoseconds), the electric force  $F_e$  is countered by a drag force  $F_d$  (i.e.  $F_e+F_d=0$ ), wherein  $F_d$  is computed from the measured velocity  $v$  and the calculated frictional coefficient  $f$   $F_d=f \times v$ , where  $f=6\pi\eta R$ ,  $\eta$ =fluid viscosity, and  $R$ =particle radius). Since velocity measurements are available at least once per cycle, the true particle forces can be continually monitored.

**Particle Stoichiometry.** In yet another embodiment, a particle monolayer in the 100  $\mu\text{m}$  field of view would contain about 7 million DNA particles, each comprising 18 hairpin units of 70 bp each. A total of about 4 million bases could be sequenced, distributed among 200 templates that are each 20 kilobases in length (i.e.  $200 \times 20,000=4,000,000$ ). Therefore, the probability that an individual particle will be captured by polymerase and imaged is only about 4/7. Since each particle would carry nearly 300 linked NTPs, there is clearly a plentiful supply of nucleotides for sequencing.

## VI. Field-Switch Chamber

In another aspect, the present invention provides a sequencing chamber. The sequencing chamber has a bottom and a top. In one aspect, the sequencing chamber has a first electrode such as on one side of the chamber (e.g., the bottom) of the chamber and a second electrode such as on another side (e.g., the top) of the chamber. A spacer separates the first electrode and second electrode. The spacer can be a polymer, an insulator, a dielectric, a semiconductor or a combination thereof. In yet another aspect, the chamber includes at least one more electrode (e.g., reference electrode, such as Ag/AgCl). The electrodes can be any geometry, such as a ring, a circle, a square, a triangle, and the like.

As shown in FIG. 7, in one embodiment, the sequencing chamber comprises two working surfaces (e.g., a bottom electrode to which polymerases are immobilized and a top electrode) where particles accumulate as they transit back and forth during the field-switch cycle. The two electrode surfaces are separated by about 5 to 10 microns. In certain aspects, the field-switch cycle operates within the constraints of polymerase kinetics, with the duration of the catalytic step being from about 1-100 msec. Suitable working surfaces for the top electrode include, without limitation, a titanium surface coated with a non-stick permeation layer.

In certain aspects, the first electrode and the second electrode independently comprise any conductor such as a metal, a metal silicide, carbon or a metal oxide. Suitable metals include, but are not limited to, indium, tin, indium-tin, silver, platinum, palladium, titanium, aluminum and gold. Suitable metal oxides include, but are not limited to, indium oxide, tin oxide, indium-tin oxide (ITO), titanium oxide, silver oxide, and aluminum oxide. In certain preferred aspects, the spacer comprises PDMS, SU-8, PET or a combination thereof. In certain other aspects, the first electrode is ITO and is coated with a permeation layer. The second electrode may also have a permeation layer. The permeation layer can be fabricated by controlled polymerization, electropolymerization adsorption, deposition (e.g., from plasma) or by self assembly.

In one embodiment, an electrode permeation layer protects the polymerase, DNA, and particles from electrochemical reactions at the electrode surface, while allowing access to ions and water. In another embodiment, the permeation

layer is used to avoid for example, the ~5 nm electrical double layer at the electrode surface, where most of the potential drop occurs between electrodes (Bard et al., In "Electrochemical Methods: Fundamentals and Applications," Wiley & Sons (2001)). Beyond the double layer, the electric field is essentially constant and particles will experience an essentially constant force. The present invention uses a thinner 5-10 nm layer in order to position the polymerases in the evanescent illumination zone. Permeation layers can be formed on ITO by self-assembled monolayers of undecanoic acid (Liron et al., *Biosens. Bioelectron.*, 17:489 (2002)). Alternatively, organic polymers can be deposited in thin (<8 nm) layers by electropolymerization; the polymer itself is nonconducting, but ion-permeable, and it provides a support for covalent protein attachment (Curulli et al., *Electroanalysis*, 13:236 (2001)). Other alternative surface coatings include electropolymerization of biotin-pyrroles, adsorption of (DOPA)3-mPEG, anodization with diols, and crown ether plasmas (respectively: Cosnier et al., *Anal. Chem.* 71:3692 (1999), Dalsin et al., *Langmuir* 21:640 (2005), Guo et al., *Chem. Pharm. bull* 44:860 (1996), Denes et al., *J Appl Polymer Sci* 81:3425 (2001)). These surfaces can be further passivated with polymers or proteins or other polymers such as polyethylene oxide (PEG), bovine serum albumin, streptavidin, avidin or neutravidin to reduce non-specific adsorption. In one embodiment, streptavidin and related proteins also provide for specific oriented attachment of biotinylated proteins and polymerases. The permeation layer can also be made from a material such as a lipid, which simulates a membrane bound protein.

In one embodiment, the surface is a metal-oxide film such as indium-tin oxide (ITO) or platinum-silicide. In another embodiment, the surface is a metal film such as gold, silver or palladium. In yet another embodiment, the electrode(s) is coated by a permeation layer to isolate immobilized complexes both from electrochemical reactions and from the electrical double-layer occurring at the electrode surface. In other embodiments, the permeation layer thickness is chosen to minimize fluorescence quenching by metals in the electrode while maintaining the complexes in the evanescent wave associated with total internal reflection microscopy. In still other embodiments, the angle of incidence of the excitation light is chosen to allow for surface plasmon excitation at the electrode surface. In certain instances, the plurality of complexes are immobilized onto the surface by covalent bonding. In certain other instances, the plurality of complexes are immobilized onto the surface by ionic bonding. In still other instances, the plurality of complexes are immobilized onto the surface by macromolecular bonds such as avidin-biotin or antibody-epitope interactions.

In another embodiment, the top electrode is, for example, a titanium or gold disk covered with a non-stick permeation layer as described for the bottom electrode (above).

## VII. EXAMPLES

The following examples are offered to illustrate, but not to limit, the claimed invention.

### Example 1. Field-Switch Chamber Construction

This example illustrates the construction of a 10  $\mu\text{m}$   $\times$  250  $\mu\text{m}$  two-electrode sequencing chamber of the present invention comprising a spin-coated SU-8 spacer permanently bonded to an ITO-coated coverslip. The ITO surface in the bottom of the chamber is then coated with a permeation

layer by controlled electropolymerization with biotin-pyrrole. The chamber is closed with a top electrode.

Chamber Construction. Microscope coverslips (25 $\times$ 25 $\times$ 0.17 mm) are coated with ITO to a resistivity of 140 ohm/square and refractive index 1.52 (ZC&R Inc.). The coverslips are cleaned by successive washes in Decon 77, deionized water and chloroform. After oven-drying for 30 min and cooling to room temperature, the coverslips are coated with 10  $\mu\text{m}$  of SU-8 photoresist by spin casting. The resist is exposed with collimated UV light through a contact mask to give a 250  $\mu\text{m}$  diameter hole and a 2 $\times$ 2 mm contact pad. The photoresist is then baked and developed (Shipley Microposit EC solvent).

ITO Permeation Layer. The ITO surface in the bottom of the 250  $\mu\text{m}$   $\times$  10  $\mu\text{m}$  PDMS well is next coated with a permeation layer by electropolymerization with biotin-pyrrole as described in Kuramitz et al., *Electroanalysis* 15:225 (2003). The polymer surface is then coated with Neutravidin (Konry et al., *Anal Chem* 75:2633 (2003)).

Top Electrode. A titanium disk (3 mm diameter $\times$ 2 mm thick) coated with mPEG-DOPA<sub>3</sub> (Dalsin et al., *Langmuir* 21:640 (2005)) is used as the top electrode. After a sample is applied to the chamber, excess fluid is squeezed out as the top electrode is set in place and the chamber is sealed by a drop of mineral oil. Electrical contact is made to the top exposed surface of the electrode disk.

### Example 2. Construction of a DNA Nanoparticle By Rolling Circle Replication

Template construction. The circular template used for the rolling circle amplification was constructed by ligating two 5' phosphate-labeled oligos (oligo 1 and oligo 2) that contain regions of inverted repeated DNA, forming a hairpin structure (FIGS. 8A-B). Ten pmoles of each oligo were combined with 1 $\times$  final concentration of T4 ligation buffer (New England Biolabs) in a volume of 20  $\mu\text{L}$ . The mixture was incubated at 90 $^{\circ}$  C. for 1 min and then cooled to 20 $^{\circ}$  C. T4 DNA ligase (100 units, New England Biolabs) was then added to the reaction mixture and incubated at 24 $^{\circ}$  C. for 3 hours. The ligated template was purified on a 6% TBE-Urea denaturing acrylamide gel, cutting the ligated circular template from the gel, and extracting the DNA from the gel fragment by incubating in 10 mM Tris pH 8 for 8 hours at 20 $^{\circ}$  C. (FIG. 9A).

In one embodiment, a DNA particle is constructed by rolling circle amplification of a DNA template (FIG. 8A). This produces a DNA particle comprising multiple identical hairpin units linked along a single strand. In one embodiment, the template for the rolling circle amplification may be a primed single stranded circular DNA strand that contains a region where a specific DNA oligo primer will hybridize to the template and initiate rolling circle amplification. In another embodiment, the template is partially self-complementary, forming a "dumbbell" structure with a primer binding region on one of the single-stranded loops. In yet another embodiment, the template sequence is constructed in such a way to allow incorporation of a modified nucleotides, such as 5-aminohexylacrylamido-dUTP (Aha-dUTP) (Molecular Probes), at specific positions along the amplified product (FIG. 8A-B). Rolling circle amplification can be accomplished either with DNA or RNA polymerases having strand displacement activity. In one embodiment, the template may be designed to provide restriction enzyme recognition sites in order to trim the amplified product to a desired size range, giving particles comprising one or more units. In another embodiment, the particle size can be controlled in the rolling

circle amplification reaction by limiting the reaction time or limiting the amount of dNTPs.

Rolling circle amplification. A mixture of 30 fmoles of template DNA, 500 fmoles of an oligo primer, 1 mM final concentration of each natural or modified dNTP, 1x final concentration of phi29 reaction buffer (New England Biolabs), 100 µg/ml BSA, and 5 units of phi29 DNA polymerase were incubated in a 25 µL volume for 3 hours at 30° C. The amplified product consists of repeating units of DNA hairpins (FIG. 8C)(FIG. 9B). mFold computer analysis, estimating the most stable confirmation for the DNA structures, for a single hairpin unit and for a series of six hairpin units is shown (FIG. 8D).

Enzymatic Cleavage of rolling circle amplified products. Rolling circle products made with dATP, dCTP, dGTP, and either dTTP (FIG. 10B lanes 3, 5, 7, and 9) or Aha-dUTP (FIG. 9B lanes 4, 6, 8, and 10) were produced. The products were digested with restriction enzymes BstUI (recognition site CGCG) BstUI (FIG. 9B lanes 5 and 6), BssHII (recognition site GCGCGC) (FIG. 9B lanes 7 and 8), or AscI (recognition site GGCGCGCC) (FIG. 9B lanes 9 and 10). Lanes 1 and 2 contain 50 bp and 500 bp DNA standard ladders.

#### Example 3. Synthesis of dCTP-PEGS-Cys-SYBR-101

Deprotection of BOC-amino PEG amine (FIG. 10, cpd 1). BOC-amino PEG amine (PolyPure, 1 gram), was added to a trifluoroacetic acid/chloroform solution (1:1 v/v, 20 mL). Reaction was stirred at room temperature for 2 hours. NaHCO<sub>3</sub> (10 mL, 10 mM) was slowly added to the reaction. The aqueous layer was separated from the organic layer and retained. Chloroform was added and two extractions (20 mL each) were performed. To the aqueous layer, ethyl ether (20 mL) was added and the extraction was performed twice. The aqueous layer was dried in vacu. The final product was a clear oily residue, and weighed approximately 1 gram. Yield was estimated at less than 100% due to possible salts and water still present in sample. <sup>1</sup>H NMR, D<sub>2</sub>O: delta 3.675 (t, 4H), 3.61 (m, 24H), 3.1 (t, 4H). ESI MS [M+1] calculated 369.3, observed [M+1] 369.2

dATP-PEGS-amine (FIG. 11, cpd 2). 13.3 mg dATP from Sigma (24.8 µmol, 1 equiv) and 190.6 mg EDC from Aldrich (992 µmol, 40 equiv) were added together in 500 mM IVIES buffer pH 5.8, giving final concentrations 79 mM and 3.2M, respectively. The reaction proceeded at room temperature for 10 minutes. A PEG-diamine (1) solution (750 mM, 186 µmol, 7.5 equiv) in 500 mM K-MES buffer pH 6.0 was added to the dATP/EDC (248 µL). The reaction proceeded at room temperature overnight. The large excess of reagents and overnight reaction converted the entire amount of starting material to desired product. The product was first purified on a HiPrep™ DEAE 20 mL column from Amersham with buffers A and B (A: 10 mM phosphate+20% ACN and buffer B: Buffer A in 1M NaCl). LC conditions: 0-10 min 0% B, 10-15 min 0-100% B, 15-20 min 100% B at a flow rate of 10 mL/min. The free PEG eluted from the column in void volume. The nucleotide conjugate product eluted as a broad peak at 5-10 minutes. This broad peak was collected and dried in vacu. The product was further purified on a 30x250 mm 10 µm Inertsil ODS-3 column from Varian, with buffer system A and B (A: 100 mM TEAAc (pH 6.6) with 4% ACN and B: 100 mM TEAAc (pH 6.6) with 80% acetonitrile). LC conditions: 0-5 min 0-15% B, 5-10 min 15-20% B, 10-20 min 20-100% B with a flow rate of 20 mL/min. The product eluting at 12 minutes was collected

and dried in vacu. Yield averaged 30% after purification steps. <sup>1</sup>H NMR D<sub>2</sub>O: delta 8.52 (s, 1H), 8.27 (s, 1H), 6.5 (t, 1H), 4.3 (t, 1H), 4.2 (m, 1H), 4.1 (m, 1H), 3.8 (m, 29H overlaid with contaminant), 3.2 (t, 2H+contaminant), 3.1 (m, 2H+contaminant), 2.8 (m, 1H), 2.6 (m, 1H). Contaminants present in sample were TEAAc and glycerol. ESI-MS [M+1] calculated 842.25, observed [M+1] 842.3

SYBR101-Cysteine (FIG. 12, cpd 3). SYBR101-NHSE purchased from Molecular Probes (1 mg/mL in DMF, 1.7 mM, 400 µL, 6.8e<sup>-4</sup> mmol) was added to a cysteine solution dissolved at 50 mM with carbonate buffer pH adjusted to 8.3 (Sigma, 100 mM, 6.8 µL, 6.8e<sup>-4</sup> mmol). The two compounds were added together and allowed to react at room temperature, protected from light, for 30 minutes. ESI-MS[M+1] calculated 585.2, observed [M+1] 585.2. The product was not purified. Estimated product yield: 80%.

dCTP-PEG-IA (FIG. 13, cpd 4). This reaction should be run at the same time as the previous reaction. dCTP-PEGS-amine (FIG. 11, cpd 2; dCTP starting nucleotide instead of dATP) was reconstituted in DMSO (20 mM, 34 µL, 6.8e<sup>-4</sup> mmol). SIA purchased from Pierce was dissolved in DMSO (10 mM, 68 µL, 6.8e<sup>-4</sup> mmol). The final concentration of nucleotide in the reaction was adjusted to 5 mM with 34 µL buffer (100 mM sodium phosphate, 0.15M NaCl, pH 7.4). Reaction proceeded at room temperature, protected from light, for 30 min. Reaction, monitored by C18 HPLC, was finished after 30 min. The dCTP-PEG-IA product was not purified. Estimated product yield: 75%.

dCTP-PEGS-cys-SYBR101 (FIG. 14, cpd 5). Entire reactions of unpurified (3) and (4) were added together. Reaction proceeded at room temperature for 18 hours, protected from light. Reaction was then purified by HPLC. LC conditions were 20-100% B over a period of 15 min (buffer A was 100 mM TEAAc (pH 6.6) with 4% ACN and buffer B was 100 mM TEAAc (pH 6.6) with 80% acetonitrile at a flow rate of 3 mL/min. LC column was a Polaris 3 µm, 10x150 mm (Varian). Product eluted at 8 minutes and was dried in vacu. LC ESI-MS[M+1] calculated 1442.4 Observed [M+1] 1442.4. Estimated product yield: 80%.

#### Example 4. Illustrates the Use of Applied Force

Nanoparticle transport out of the evanescent detection zone is fast (i.e., <0.5 msec) so that bound particles can be imaged prior to completion of the base addition reaction (i.e., 1-100 msec). Particle transport can be arbitrarily fast if enough force is used. Two force constraints related to (1) stretching of the DNA template and to (2) the strength of nucleotide binding to the polymerase have been identified.

The first force constraint comes from observations that physical stretching of a DNA template influences the activity of RNA and DNA polymerases (Wuite et al., *Nature*, 404: 103 (2000); Forde et al., *PNAS*, 99:11682 (2002); Goel et al., *PNAS*, 98:8485 (2001)). For T7 DNA polymerase, as the DNA stretching force increases from 0 to 35 pN, polymerase activity initially increases slightly as the force ramps up to ~5 pN, but then decreases back to the initial rate at ~10 pN, and finally is inhibited completely at a stretching force 35 pN. The slight activity increase at low force may relate to the straightening of random-coil DNA, making it a better substrate for the enzyme, whereas the inhibition at a higher force may be caused by a change in helical pitch and DNA denaturation (Rouzina et al., *Biophys. J.*, 80:882 (2001)). As such, an upper stretch force limit of 10 pN was set based on the observation that forces above this level inhibit polymerase activity. For a 20 kb DNA template (q=-40,000), the

10 pN limit would occur at a relatively low field strength (i.e.,  $E=1600$  V/m ( $f=q \times E$ ; see also below)).

The second force constraint is related to the physical strength of the bond between a nucleotide triphosphate and a polymerase. The electric force on the nucleotide must not be so great as to pull the nucleotide from the polymerase before the catalytic reaction is completed. For example, T7 DNA polymerase in a closed conformation displays a strong interaction with a nucleotide, whereby the ribose and triphosphate moieties of the nucleotide are contacted extensively by conserved residues and by two metals that coordinate the unesterified oxygens of all three phosphates (Doublie et al., *Nature*, 391:251 (1998)). Such interactions also apply to most other DNA polymerases (Patel et al., *Nat. Struct. Biol.*, 8:656 (2001)). The catalytic reaction occurs in the closed conformation, in which the nucleotide is tightly bound. After nucleotide incorporation, the polymerase changes to an open conformation from which pyrophosphate is released.

To estimate the force limit in the systems of the present invention, a simple and conservative approach was taken to first estimate the thermal force scale likely to produce rapid, exponential increases in nucleotide dissociation rate and then limit the applied force to lower values. This approach ensures that the applied electrical force is weak compared to the level of force needed to rapidly dissociate the nucleotide and the nucleotide will remain bound under force up to its natural time for spontaneous dissociation. To make the

gen bonds,  $x_\beta \sim 3-5$  Å. Since  $k_B T = 4.1$  pN-nm at room temperature, this conservative estimate leads to a much smaller value of thermal force scale,  $f_\beta = k_B T / x_\beta = 8-13$  pN. For the systems of the present invention, an upper limit of 1 pN on the particle, less than 12% of thermal force, was arbitrarily set. A 1 pN force should have negligible effect on a bound nucleotide. The required electric field strength depends on particle charge, which in turn depends on particle structure.

#### Example 5 Illustrates a Computer Model of the System

A computer model was developed to simulate individual particles under electric and thermal forces (random-walk diffusion) while tracking the fluorescence signal over time. In an evaluation of nearly 7,000 combinations of different values of the 8 input parameters, the best performing particles were small and highly charged, allowing them to experience enough force ( $f=q \times E$ ) for fast transport, given the two force constraints described above. The electrokinetic properties of 5 example particles are shown in FIG. 15 and Table III. In each case, the E-field was the highest possible, limited either by force on the particle (1 pN) or force on the template (10 pN). The total fluorescence signal ( $S_T$ ) detected at one pixel of a CCD camera is the sum of contributions ( $s_i$ ) from all particles in that pixel ( $S_T = \sum s_i$ ). The signal is scaled so that  $s_i = 1$  for an individual particle excited by the maximum evanescent energy at the electrode surface.

Table III. Particle Transport Kinetic Parameters.

Parameter	dNT-Dye	Dendrimer	DNA Circ	DNA Circ	SiDNA Part
Bases	1 base	18 kb	4 kb	2 kb	3 kb
Dia, nm	3	180	300	200	55
Charge q	-5	-3600	-7200	-3600	-5400
Diff m <sup>2</sup> /s	1.5E-09	4.8E-12	1E-11	1.6E-11	1.6E-11
Fric kg/s	3.1E-12	9.4E-10	4.5E-10	2.9E-10	2.8E-10
E-field V/m	1560	168	867	1560	1160
F(particle) pN	0.0012	1	1	0.9	1
F(DNA) pN	10	1.1	5.56	10	7.7
Veloc nm/usec	0.4	0.47	2.2	3.1	3.5
td usec	>1000	>1000	360	260	218
Dye Content	1	400	400	200	300
Illum Index	1	0.67	0.25	0.44	0.85
Dye Equiv	1	268	100	88	255

calculation, a thermal force scale  $f_\beta$  can be determined from the ratio of thermal energy  $k_B T$  (Newton-meter) to the distance  $x_\beta$  (meter) expected to be gained along the force direction in nucleotide dissociation according to the formula:  $f_\beta = k_B T / x_\beta$ . Of particular relevance here, the length scale  $x_\beta$  has been measured for the weak interaction between a protein, L-selectin, and its polysaccharide ligand, sLe<sup>x</sup> (Evans et al., *PNAS*, 98:3784 (2001)). Two length scales were found for both a strong interaction with a length  $x_\beta = 0.6-1.0$  Å (likely due to coordination of sLe<sup>x</sup> by one Ca<sup>++</sup> atom) and a weak interaction with length  $x_\beta = 4$  Å (likely due to electrostatic and hydrogen bonds).

The L-selectin example is similar to polymerase-nucleotide binding in the sense that both involve metal ion coordination and both have a like multiplicity of weaker electrostatic and hydrogen bonds. This comparison suggests that the nucleotide, coordinated by two Mg<sup>++</sup>, is likely to be bound as tightly as the fucose ring of the polysaccharide, which is coordinated by one Ca<sup>++</sup>. Although the small length scale for metal ion coordination implies a very large thermal force scale, conservative estimations can be made such that the bond rupture distance is governed by the weaker hydro-

In FIG. 15,  $S_T$  was monitored for 1 msec as a monolayer of particles moved outward from the surface. The initial signal  $STO=13$  for a monolayer of 55 nm particles. This is because roughly 13 particles comprise a monolayer over a  $200 \times 200$  nm pixel area ( $(200)^2 / (55)^2 = 13$ ). In order to compare signals across all 5 particle types, 13 particles were modeled arbitrarily for the other 4 types, regardless of their diameter. A time  $t_d$  is defined when  $S_T$  falls below a threshold  $S_T < 0.1$ ; that is, after time  $t_d$ , an individual bound particle  $s_i = 1$  can be detected with little background from unbound particles. FIG. 15 shows good results with low mass DNA structures:  $t_d = 360$  and  $260$  μsec for 4 kb and 2 kb circular DNA (curves 3 and 4, respectively); and  $t_d = 218$  μsec for the 55 nm silica-DNA particle (curve 5). The computer model indicates that fast transport kinetics are possible.

#### A. Nanoscale Particle Size to Maximize Detectability and Nucleotide Concentration

Detectability. An illumination index was computed for each particle type (see, Table III above) by summing the position-weighted illumination of each dye location and dividing by the total number of dyes in the particle. As shown in Table III, a silica-DNA particle (300 dyes) would

be more than twice the brightness of the 4 kb DNA circle (400 dyes), even though the small particle contains fewer dyes. As such, small particles are preferred because they are more fully illuminated.

Nucleotide Concentration. The field-switch sequencing method of the present invention offers the advantage of actively concentrating particles at the electrode surface, providing a high concentration of dNTPs for the immobilized polymerases during the binding phase of the sequencing cycle. Small particles are preferred because they have a higher diffusivity (i.e., more collisions with immobilized polymerases), a higher surface-to-volume ratio (i.e., more dNTPs attached per unit particle volume), and superior coverage (i.e., many small particles cover the electrode surface better than a few large particles).

$K_m$  values for free nucleotides in solution binding to DNA polymerases are typically  $<0.1$  mM. To help estimate the binding ability of DNA polymerases for nucleotides linked to nanoparticles, the density of nucleotides was calculated on a particle surface (nucleotides/area). The value obtained was then correlated to bulk molar nucleotide concentrations (nucleotides/volume). The calculation involved computing the nucleotide parking area  $pA$  (i.e.,  $\text{\AA}^2/\text{nucleotide}$ ), wherein  $pA = (\text{particle surface area})/(\text{nucleotides per particle})$ . If the surface-bound nucleotides are thought of as being in bulk solution, each would occupy a parking volume  $pV$  (i.e.,  $\text{\AA}^3/\text{nucleotide}$ ), such that  $pV = pA^{3/2}$ . In this view, 100 surface nucleotides on a 55 nm particle equate to 1.8 mM, which is sufficient to support fast nucleotide binding based on DNA polymerase binding kinetics in free solution (i.e.,  $K_m < 0.1$  mM).

While high nucleotide surface density, flexible tethered nucleotides, and high particle concentrations at the electrode can all have positive effects on polymerase binding kinetics, the lower diffusivity of particles compared to free nucleotides can have a negative effect. However, every polymerase does not have to bind a nucleotide every cycle. Missed binding events can be remedied in subsequent cycles, with the only impact being on system efficiency.

#### B. Fast Detection in Four Colors

Detection. In one embodiment, a four color single molecule imaging microscope with total internal reflectance (evanescent wave) excitation and a CCD camera to image a  $100 \mu\text{m} \times 100 \mu\text{m}$  field of view can be used. The camera is, e.g., a Micromax 512x512 or a Pentamax intensified CCD from Roper Scientific. However, the Micromax is better with respect to the signal to noise (SN). The mean background noise with the laser on is 10 electrons. With  $2 \times 2$  pixel binning and a 2 MHz readout, the total frame readout time is 33 msec, which is compatible with a 100 msec field-

switch sequencing cycle. Single immobilized TAMRA molecules are typically imaged in 80 msec with an SN of  $\sim 11$ .

Since SN scales linearly with integration time, TAMRA would be undetectable ( $SN < 0.28$ ) in the 1-2 msec timeframe required for field-switch sequencing. Although photodiode point detectors have greater sensitivity than CCD detectors, one objective of the present invention is to increase throughput 200-300 fold over what a point detector could deliver by imaging many single, immobilized polymerase-DNA complexes simultaneously with a camera.

Stability of Intercalated Complex. The NMR structure of a TOTO-DNA intercalation complex reveals a distorted double helix unwound by  $12^\circ$  and extended in length about 12% (Spielmann et al., *Biochemistry*, 34:8542 (1995)). Nonetheless, dye intercalation stabilizes DNA against denaturation, as shown by a  $T_m$  change from  $65^\circ$  C. to  $95^\circ$  C. upon intercalating 6 YOYO dyes into a 30-mer dsDNA oligonucleotide (Bjorndal et al., *Biopolymers*, 65:40 (2002)). This dye family is marketed for staining DNA prior to electrophoresis based on experiments showing that complexes comprising 1 dye:7-10 base pairs remain fluorescent in agarose gels (Rye et al., *Nucleic Acids Res.*, 20:2803 (1992)). These findings have been extended by quantifying the JOJO:DNA ratio of complexes in a dialysis bag in an electric field (1000 V/meter). At different times, the sample was removed from the bag and the dye:DNA ratio was determined spectrophotometrically. The sample was returned to the same bag and the process was repeated for the next time point. Variations in sample recovery at each time point had no effect on the measured dye:DNA ratio. In this experiment, the dye:DNA ratio was stable. In another embodiment, dyes such as SYBR-101, having a functional group for conjugation, would not only intercalate into DNA but would also be covalently linked to amino-modified DNA, are used in the present invention. Covalent linkage can thus prevent any possibility of dissociation of the dye from the DNA under conditions of elevated temperature, salt concentration, detergents or organic solvents that tend to decrease the binding affinity between dye and DNA.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

---

#### SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 6

<210> SEQ ID NO 1

<211> LENGTH: 23

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic biotin loop

<220> FEATURE:

<221> NAME/KEY: MOD\_RES

<222> LOCATION: (13) . . . (13)

<223> OTHER INFORMATION: Lys modified by biotin

<400> SEQUENCE: 1



-continued

---

Ser Ser Gly Gly Leu Asn Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp  
 1 5 10 15

His Glu Asp Thr Gly Ser Ser  
 20

<210> SEQ ID NO 2  
 <211> LENGTH: 85  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: synthetic oligonucleotide oligo 1 for template construction

<400> SEQUENCE: 2

cagcgtgccg cgctgggacg gtgcgagcct tggtaggct cgcaccgtcc cagcgtggca 60  
 cgctgcgacc ctgcgacggt gccac 85

<210> SEQ ID NO 3  
 <211> LENGTH: 86  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: synthetic oligonucleotide oligo 2 for template construction

<400> SEQUENCE: 3

gctcgcaggc tggcgcgccc ttccggatcg ggtagggtc ctatttcggc gcgccagcct 60  
 gcgagcgtgg caccgtcgca gggtcg 86

<210> SEQ ID NO 4  
 <211> LENGTH: 20  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: synthetic primer 1 for rolling circle replication

<400> SEQUENCE: 4

taggacccta acccgatccg 20

<210> SEQ ID NO 5  
 <211> LENGTH: 164  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: synthetic single hairpin unit

<400> SEQUENCE: 5

tttcggcgcg cctgcgagcg tggcaccgtc gcagggtcgc agcgtgccac gctgggacgg 60  
 tgcgagcctt ggtaggctc gcaccgtccc agcgtggcac gctgcgaccc tgcgacgggtg 120  
 ccacgctcgc aggcgcgccc tttgggatcg ggtagggtc ctaa 164

<210> SEQ ID NO 6  
 <211> LENGTH: 21  
 <212> TYPE: PRT  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: synthetic 9 Degrees North DNA polymerase anchor  
 <220> FEATURE:  
 <221> NAME/KEY: MOD\_RES  
 <222> LOCATION: (11)...(11)  
 <223> OTHER INFORMATION: Xaa = p-acetyl-L-phenylalanine (pa-Phe)

<400> SEQUENCE: 6

-continued

---

Leu Leu Ser Lys Lys Arg Ser Leu Cys Cys Xaa Cys Thr Val Ile Val  
 1 5 10 15

Tyr Val Thr Asp Thr  
 20

---

What is claimed is:

1. A polymerase-nucleic acid complex attached to a solid support, the polymerase-nucleic acid complex comprising:

a target nucleic acid, a polymerase comprising a nucleic acid binding cleft, and a topological tether, wherein the topological tether is attached to the polymerase through at least two anchors that straddle the nucleic acid binding cleft, whereby the topological tether entraps the target nucleic acid in the nucleic acid binding cleft of the polymerase, and wherein the polymerase-nucleic acid complex is attached to a solid support.

2. The polymerase-nucleic acid complex of claim 1, wherein at least one of the anchors comprises at least one amino acid of the polymerase or an epitope of the polymerase.

3. The polymerase-nucleic acid complex of claim 2, wherein the at least one amino acid is selected from the group consisting of a cysteine, a phenylalanine derivative and a histidine.

4. The polymerase-nucleic acid complex of claim 3, wherein the at least one amino acid is a cysteine.

5. The polymerase-nucleic acid complex of claim 1, wherein the nucleic acid binding cleft is a DNA binding cleft.

6. The polymerase-nucleic acid complex of claim 1, wherein the at least two anchors comprise cysteines of the polymerase.

7. The polymerase-nucleic acid complex of claim 1, wherein the at least two anchors each comprise a first member of a complementary binding pair that is bound to a second member of the complementary binding pair, which second member is located on the topological tether, whereby the topological tether is attached to the polymerase.

8. The polymerase-nucleic acid complex of claim 7, wherein the binding pair is biotin-streptavidin or biotin-avidin.

9. The polymerase-nucleic acid complex of claim 1, wherein the topological tether is covalently attached to the polymerase.

10. The polymerase-nucleic acid complex of claim 1, wherein the topological tether irreversibly associates the target nucleic acid with the polymerase.

11. The polymerase-nucleic acid complex of claim 1, wherein the topological tether is directly anchored to the solid support.

12. The polymerase-nucleic acid complex of claim 1, wherein the polymerase-nucleic acid complex further comprises a primer nucleic acid which complements a region of the target nucleic acid.

13. The polymerase-nucleic acid complex of claim 1, wherein the target nucleic acid is a circular DNA.

14. The polymerase-nucleic acid complex of claim 1, wherein the polymerase is selected from the group consisting of Klenow, Taq, and T7 polymerase.

15. The polymerase-nucleic acid complex of claim 1, wherein the polymerase is selected from the group consisting of phi29, RB-69 and T4 polymerase.

16. The polymerase-nucleic acid complex of claim 15, wherein the polymerase is phi29.

17. The polymerase-nucleic acid complex of claim 1, wherein the polymerase is a reverse transcriptase.

18. The polymerase-nucleic acid complex of claim 1, wherein the polymerase is an HIV reverse transcriptase.

19. The polymerase-nucleic acid complex of claim 1, wherein the polymerase is a DNA polymerase from *Thermus flavus*, *Pyrococcus furiosus*, *Thermotoga neapolitana*, *Thermococcus litoralis*, *Sulfolobus solfataricus*, *Thermatoga maritima*, or *E. coli*.

20. A composition comprising the polymerase-nucleic acid complex of claim 1 and detectably labeled nucleoside triphosphates.

21. A polymerase-nucleic acid complex, the polymerase-nucleic acid complex comprising:

a circular target nucleic acid, a polymerase comprising a nucleic acid binding cleft, and a topological tether, wherein the topological tether is covalently attached to the polymerase through at least two anchors that straddle the nucleic acid binding cleft, thereby irreversibly associating the circular target nucleic acid with the nucleic acid binding cleft of the polymerase; and wherein said polymerase-nucleic acid complex shows a higher processivity index than a polymerase-nucleic acid complex without said topological tether.

22. A polymerase-nucleic acid complex, the polymerase-nucleic acid complex comprising:

a target nucleic acid, a polymerase comprising a nucleic acid binding cleft, and a topological tether, wherein the topological tether is attached to the polymerase through at least two anchors that straddle the nucleic acid binding cleft, whereby the topological tether entraps the target nucleic acid in the nucleic acid binding cleft of the polymerase, and wherein the polymerase-nucleic acid complex further comprises a DNA primer which complements a region of the target nucleic acid.

23. A polymerase-nucleic acid complex, the polymerase-nucleic acid complex comprising:

a target nucleic acid, a polymerase comprising a nucleic acid binding cleft, and a topological tether, wherein the topological tether is attached to the polymerase through at least two anchors that straddle the nucleic acid binding cleft, whereby the topological tether entraps the target nucleic acid in the nucleic acid binding cleft of the polymerase, and wherein the target nucleic acid is a circular DNA.

24. A polymerase-nucleic acid complex, the polymerase-nucleic acid complex comprising:

a target nucleic acid, a phi29 polymerase comprising a nucleic acid binding cleft, and a topological tether, wherein the topological tether is attached to the phi29 polymerase through at least two anchors that straddle the nucleic acid binding cleft, whereby the topological tether entraps the target nucleic acid in the nucleic acid binding cleft of the phi29 polymerase.

25. A polymerase-nucleic acid complex, the polymerase-nucleic acid complex comprising:

a target nucleic acid, a reverse transcriptase comprising a nucleic acid binding cleft, and a topological tether, wherein the topological tether is attached to the reverse transcriptase through at least two anchors that straddle the nucleic acid binding cleft, whereby the topological tether entraps the target nucleic acid in the nucleic acid binding cleft of the reverse transcriptase. 5

**26.** A composition comprising a polymerase-nucleic acid complex and detectably labeled nucleoside triphosphates, the polymerase-nucleic acid complex comprising: 10

a target nucleic acid, a polymerase comprising a nucleic acid binding cleft, and a topological tether, wherein the topological tether is attached to the polymerase through at least two anchors that straddle the nucleic acid binding cleft, whereby the topological tether entraps the target nucleic acid in the nucleic acid binding cleft of the polymerase. 15

\* \* \* \* \*